LINEAR PRECODING AND DFE EQUALIZATION ACHIEVE THE DIVERSITY VS MULTIPLEXING OPTIMAL TRADEOFF

Abdelkader Medles, Dirk T.M. Slock

Eurecom Institute
2229 route des Crêtes, B.P. 193
06904 Sophia Antipolis Cedex, FRANCE
Email: {abdelkader.medles,dirk.slock}@eurecom.fr

ABSTRACT

The use of multiple transmit (TX) and receive (RX) antennas allows to transmit multiple signal streams in parallel and hence to increase communication capacity. We have previously introduced simple convolutive linear precoding schemes that spread transmitted symbols in time and space, involving spatial spreading, delay diversity and possibly temporal spreading. In this paper we show that the use of the classical MIMO DFE Equalizer for this system allows to achieve the optimal diversity versus multiplexing tradeoff introduced in [1].

1. ITRODUCTION

The $N_{tx} \times N_{rx}$ MIMO system is essentially described by

$$\boldsymbol{y}_k = \mathbf{H}\,\boldsymbol{a}_k + \boldsymbol{v}_k = \mathbf{H}\,\mathbf{T}(q)\,\boldsymbol{b}_k + \boldsymbol{v}_k \tag{1}$$

where the white noise power spectral density matrix is $S_{\boldsymbol{vv}}(z) = \sigma_v^2\, I$, and $q^{-1}\,\boldsymbol{b}_k = \boldsymbol{b}_{k-1}$. We consider the case of channel state information being absent at the transmitter (TX) and perfect at the receiver (RX). The linear precoding considered here (introduced in [2] and further analyzed in [3]) consists of a modification of VBLAST, obtained by inserting a square matrix prefilter $\mathbf{T}(z)$ before inputting the vector signal $\boldsymbol{b}_k$ into the channel $\mathbf{H}$. The N_{tx} signal components of $\boldsymbol{b}_k$ are called streams or layers. The suggested prefilter is $\mathbf{T}(z) = \mathbf{D}(z)\mathbf{Q}$ where $\mathbf{D}(z) = \text{diag}\,\{1,\ z^{-1}, \ldots, z^{-(N_{tx}-1)}\}$. $\mathbf{Q}$ is unitary $\mathbf{Q}^H\mathbf{Q} = I$:

$$\mathbf{Q} = \frac{1}{\sqrt{N_{tx}}}\begin{bmatrix} 1 & \theta_1 & \ldots & {\theta_1}^{N_{tx}-1} \\ 1 & \theta_2 & \ldots & {\theta_2}^{N_{tx}-1} \\ \vdots & \vdots & & \vdots \\ 1 & \theta_{N_{tx}} & \ldots & {\theta_{N_{tx}}}^{N_{tx}-1} \end{bmatrix}, \tag{2}$$

where the θ_i are the roots of $\theta^{N_{tx}} - j = 0$, $j = \sqrt{-1}$.
Every symbol stream m ($b_{m,k}$) passes through the equivalent SIMO channel $\sum_{i=1}^{N_{tx}} z^{-(i-1)}\mathbf{H}_{:,i}\mathbf{Q}_{i,m}$ which now has memory due to the delay diversity introduced by $\mathbf{D}(z)$. It is important that the different columns $\mathbf{H}_{:,i}$ of the channel matrix get spread out in time to get full diversity (otherwise the streams just pass through a linear combination of the columns, as in VBLAST, which offers limited diversity). The delay diversity only becomes effective by the introduction of the spatial spreading matrix $\mathbf{Q}$, which has equal magnitude elements for uniform diversity spreading (a specific choice for $\mathbf{Q}$ exists for maximum coding gain in case of QAM symbols [3]). We can see that each symbol stream has the same Matched Filter Bound (MFB), which is proportional to the channel Frobenius norm, hence full diversity is exploited. Also, since the prefilter $\mathbf{T}(z)$ is paraunitary and transforms the white stream $\boldsymbol{b}_k$ into the white stream $\boldsymbol{a}_k$, no loss in ergodic capacity is incurred. In what follows we denote the overall channel by $\mathbf{G}(z) \triangleq \mathbf{HT}(z)$.

Institut Eurécom's research is partially supported by its industrial members: Bouygues Télécom, Fondation d'entreprise Groupe Cegetel, Fondation Hasler, France Télécom , Hitachi, ST Microelectronics, Swisscom, Texas Instruments, Thales.

2. CONVENTIONAL MIMO DFE RECEIVER

Consider the classical MIMO decision feedback equalizer, in which the symbol vectors $\mathbf{b}_k$ are processed sequentially in time (see Fig. 1). The output of the matched filter is $\mathbf{G}(z)$ is $\mathbf{z}_k = \mathbf{G}^\dagger(q)\mathbf{y}_k$.

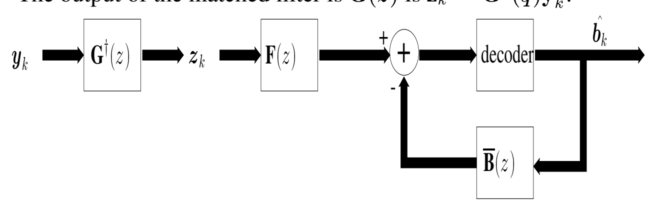

Fig. 1. MIMO DFE receiver

The DFE output is then

$$\widehat{\mathbf{b}}_k \;=\; -\underbrace{\overline{\mathbf{B}}(q)}_{\text{feedback}}\;\mathbf{b}_k + \underbrace{\mathbf{F}(q)}_{\text{feedforward}}\;\mathbf{z}_k\,, \tag{3}$$

where the feedback filter $\overline{\mathbf{B}}(z) = \sum_{i\geq 1} \mathbf{B}_i z^{-i}$ is such that $\mathbf{B}(z) = I + \overline{\mathbf{B}}(z)$ is causal, monic and minimum phase. Different designs of Rx are possible (MMSE, MMSE ZF ...), we consider here the MMSE design.

2.1. MMSE Conventional MIMO DFE Rx

The MMSE linear symbol vector estimate (MMSE linear equalizer output) verifies

$$\begin{aligned} \widehat{\mathbf{b}}_k^{mmse} &= \mathbf{S}_{\mathbf{by}}(q)\,\mathbf{S}^{-1}_{\mathbf{yy}}(q)\,\mathbf{y}_k \\ &= \sigma_a^2\mathbf{G}^\dagger(q)(\sigma_a^2\mathbf{G}(q)\mathbf{G}^\dagger(q) + \sigma_v^2\mathbf{I})^{-1}\,\mathbf{y}_k \\ &= \rho(\rho\mathbf{G}^\dagger(q)\mathbf{G}(q) + \mathbf{I})^{-1}\mathbf{G}^\dagger(q)\,\mathbf{y}_k \\ &= \mathbf{R}^{-1}(q)\,\mathbf{z}_k \end{aligned} \tag{4}$$

where $\mathbf{R}(z)=\mathbf{G}^{\dagger}(z)\mathbf{G}(z)+\frac{1}{\rho}I$ and $\rho=\frac{\sigma_b^2}{\sigma_v^2}=\frac{\sigma_a^2}{\sigma_v^2}$.
$\widehat{\mathbf{b}}_k^{mmse}$ can be also written as $\widehat{\mathbf{b}}_k^{mmse}=\mathbf{b}_k+\widetilde{\mathbf{b}}_k^{mmse}$ then

$$\mathbf{b}_k \;=\; \widehat{\mathbf{b}}_k^{mmse} - \widetilde{\mathbf{b}}_k^{mmse} \;=\; \mathbf{R}^{-1}(q)\,\mathbf{z}_k - \widetilde{\mathbf{b}}_k^{mmse}. \qquad (5)$$

Due to the orthogonality principle of the MMSE estimate we have

$$\mathbf{S}_{\widetilde{\mathbf{b}}\widetilde{\mathbf{b}}}^{mmse}(z)=\mathbf{S}_{\mathbf{bb}}(z)-\mathbf{S}_{\widehat{\mathbf{b}}\widehat{\mathbf{b}}}^{mmse}(z)=\sigma_v^2\mathbf{R}^{-1}(z). \qquad (6)$$

Consider the minimum and maximum phase factorization of $\mathbf{R}(z)$ (see [4]). Let $\mathbf{B}(z)$ be the unique causal, monic ($\mathbf{B}(\infty)=I_{N_{tx}}$) minimum phase factor of $\mathbf{R}(z)$, then

$$\mathbf{R}(z)\;=\;\mathbf{B}^{\dagger}(z)\,\mathbf{M}\,\mathbf{B}(z)\,, \qquad (7)$$

where $\mathbf{M}$ is a constant positive definite hermitian matrix.
Then $\mathbf{b}_k\;=\;\mathbf{B}^{-1}(q)\,\mathbf{M}^{-1}\,\mathbf{B}^{-\dagger}(q)\,\mathbf{z}_k-\widetilde{\mathbf{b}}_k^{mmse}$.
By choosing $\mathbf{F}(q)=\mathbf{M}^{-1}\mathbf{B}^{-\dagger}(q)$, we get

$$\begin{aligned}\mathbf{F}(q)\,\mathbf{z}_k &= \mathbf{M}^{-1}\,\mathbf{B}^{-\dagger}(q)\,\mathbf{z}_k\\ &= \mathbf{M}^{-1}\,\mathbf{B}^{-\dagger}(q)\,\mathbf{R}(q)\,(\mathbf{b}_k-\widetilde{\mathbf{b}}_k^{mmse})\\ &= \mathbf{B}(q)\,\mathbf{b}_k-\mathbf{B}(q)\,\widetilde{\mathbf{b}}_k^{mmse}\\ &= \mathbf{B}(q)\,\mathbf{b}_k+\mathbf{e}_k\\ &= \mathbf{b}_k+\overline{\mathbf{B}}(q)\,\mathbf{b}_k+\mathbf{e}_k\,,\end{aligned} \qquad (8)$$

where $\mathbf{S}_{\mathbf{ee}}(z)\;=\;\sigma_v^2\,\mathbf{B}(z)\mathbf{R}^{-1}(z)\mathbf{B}^{\dagger}(z)\;=\;\sigma_v^2\,\mathbf{M}^{-1}$.
$\overline{\mathbf{B}}(z)=\mathbf{B}(z)-\mathbf{I}$ is tightly related to the MIMO prediction error filter $\mathbf{P}(z)$ of the spectrum $\mathbf{R}(z)$, $\mathbf{P}^{\dagger}(z)\mathbf{R}(z)\mathbf{P}(z)=$ Constant Matrix. Indeed, $\mathbf{P}(z)=\mathbf{B}^{-1}(z)$ obviously. The following theorem gives $\mathbf{B}(z)$ in the case of a flat MIMO channel.

Theorem 1: For a frequency-flat MIMO channel the feedback filter is

$$\mathbf{B}(z)=\mathbf{T}(z)^{\dagger}\,\mathbf{L}^H\,\mathbf{T}(z)\,, \qquad (9)$$

with the corresponding

$$\mathbf{M}=\mathbf{Q}^H\mathbf{D}\mathbf{Q}\,, \qquad (10)$$

where $\mathbf{L}$ and $\mathbf{D}$ result from the LDU triangular matrix decomposition of $\mathbf{H}^H\mathbf{H}+\frac{1}{\rho}I=\mathbf{L}\,\mathbf{D}\,\mathbf{L}^H$.
Proof :
We need to show that $\mathbf{B}(z)\;=\;\mathbf{Q}^H\,\mathbf{D}(z)^{\dagger}\,\mathbf{L}^H\,\mathbf{D}(z)\,\mathbf{Q}$ is a minimum phase causal monic filter and verifies
$\mathbf{B}^{-\dagger}(z)\mathbf{R}(z)\mathbf{B}^{-1}(z)=\mathbf{M}$.
$\mathbf{L}^H$ is upper triangular with unit diagonal, then due to the diagonal structure of $\mathbf{D}(z)$, $\mathbf{D}(z)^{\dagger}\,\mathbf{L}^H\,\mathbf{D}(z)$ is a monic causal filter. $\mathbf{Q}$ is unitary, hence $\mathbf{B}(z)$ is also a causal monic filter.
$\det\mathbf{B}(z)\;=\;\det\mathbf{L}^H\;=\;1$, which shows that $\mathbf{B}(z)$ is minimum phase. To complete the proof of the theorem it is sufficient to verify that $\mathbf{B}^{-\dagger}(z)\mathbf{R}(z)\mathbf{B}^{-1}(z)=\mathbf{Q}^H\mathbf{D}\mathbf{Q}=\mathbf{M}$. □

2.2. Unbiased MMSE Conventional MIMO DFE Rx

$\mathbf{F}(q)\,\mathbf{z}_k-\overline{\mathbf{B}}(z)\,\mathbf{b}_k$ is a biased estimate of $\mathbf{b}_k$, since

$$\begin{aligned}\mathbf{F}(q)\,\mathbf{z}_k-\overline{\mathbf{B}}(q)\,\mathbf{b}_k &= [\mathbf{M}^{-1}\,\mathbf{B}^{-\dagger}(q)\,\mathbf{G}^{\dagger}(q)\mathbf{G}(q)-\overline{\mathbf{B}}(z)]\mathbf{b}_k\\ &\quad+\mathbf{M}^{-1}\,\mathbf{B}^{-\dagger}(q)\mathbf{G}^{\dagger}(q)\boldsymbol{v}_k\\ &= (I-\tfrac{1}{\rho}\,\mathbf{M}^{-1})\,\mathbf{b}_k+\tilde{\mathbf{e}}_k\,,\end{aligned} \qquad (11)$$

where

$$\tilde{\mathbf{e}}_k=\mathbf{M}^{-1}\,\mathbf{B}^{-\dagger}(q)\mathbf{G}^{\dagger}(q)\boldsymbol{v}_k-\frac{1}{\rho}\,\mathbf{M}^{-1}\,(\mathbf{B}^{-\dagger}(q)-I)\mathbf{b}_k. \qquad (12)$$

The covariance of $\tilde{\mathbf{e}}_k$ is

$$\begin{aligned}\mathbf{C}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}} &= \tfrac{1}{2\pi j}\oint[\mathbf{M}^{-1}(\sigma_v^2\mathbf{B}^{-\dagger}(z)\mathbf{G}^{\dagger}(z)\mathbf{G}(z)\mathbf{B}(z)\\ &\qquad+\sigma_b^2\rho^{-2}(\mathbf{B}^{-\dagger}(z)-I)(\mathbf{B}^{-1}(z)-\mathbf{I}))\mathbf{M}^{-1}]\\ &= \tfrac{1}{2\pi j}\oint[\mathbf{M}^{-1}(\sigma_v^2\mathbf{B}^{-\dagger}(z)\mathbf{G}^{\dagger}(z)\mathbf{G}(z)\mathbf{B}(z)\\ &\qquad+\sigma_v^2\rho^{-1}\mathbf{B}^{-\dagger}(z)\mathbf{B}^{-1}(z))\mathbf{M}^{-1}]-\sigma_v^2\rho^{-1}\mathbf{M}^{-2}\\ &= \tfrac{1}{2\pi j}\oint\mathbf{M}^{-1}(\sigma_v^2\mathbf{B}^{-\dagger}(z)(\mathbf{G}^{\dagger}(z)\mathbf{G}(z)+\rho^{-1}\mathbf{I})\mathbf{B}^{-1}(z))\mathbf{M}^{-1}\\ &\quad-\sigma_v^2\rho^{-1}\mathbf{M}^{-2}\\ &= \tfrac{1}{2\pi j}\oint(\mathbf{M}^{-1}\sigma_v^2\mathbf{M}\mathbf{M}^{-1})-\sigma_v^2\rho^{-1}\mathbf{M}^{-2}\\ &= \sigma_v^2\,\mathbf{M}^{-1}(I-\tfrac{1}{\rho}\,\mathbf{M}^{-1})\end{aligned} \qquad (13)$$

The feedforward UMMSE filter is

$$\mathbf{F}^U(q)=(I-\frac{1}{\rho}\,\mathbf{M}^{-1})^{-1}\,\mathbf{M}^{-1}\mathbf{B}^{-\dagger}(q)=(\mathbf{M}-\frac{1}{\rho}\,\mathbf{I})^{-1}\,\mathbf{B}^{-\dagger}(q), \qquad (14)$$

whereas the corresponding feedback filter is

$$\overline{\mathbf{B}}^U(q)=(I-\frac{1}{\rho}\,\mathbf{M}^{-1})^{-1}(\mathbf{B}(q)-\mathbf{I}). \qquad (15)$$

The output of the DFE is then

$$\begin{aligned}\widehat{\mathbf{b}}_k^U &= \mathbf{F}^U(q)\,\mathbf{z}_k-\overline{\mathbf{B}}^U(q)\,\mathbf{b}_k\\ &= \mathbf{b}_k+\tilde{\mathbf{e}}_k^U\,,\end{aligned} \qquad (16)$$

where $\mathbf{C}_{\tilde{\mathbf{e}}^U\tilde{\mathbf{e}}^U}=\sigma_v^2\,\mathbf{M}^{-1}(I-\frac{1}{\rho}\,\mathbf{M}^{-1})^{-1}$.

3. DIVERSITY VS MULTIPLEXING TRADEOFF

In [1], Zheng and Tse introduced the diversity versus multiplexing tradeoff. In what follows, we study the diversity vs multiplexing tradeoff achieved by the Conventional MIMO DFE equalizer, applied to our linearly precoded system. We consider a transmission over a large frame of length T ($T>>N_{tx}$). As the delay introduced by $\mathbf{T}(q)$ is $N_{tx}-1$, then the number of symbol vectors $\mathbf{b}_k$ transmitted over the frame duration is $T-N_{tx}+1$ (padded by $N_{tx}-1$ transmitted zeros at the end of each frame). As the considered frame size is large ($T>>N_{tx}$), we can then neglect the effect on the rate that results for the loss of $N_{tx}-1$ symbol periods.

Theorem 2: In the case of a frequency-flat channel and $N_{tx}=2^n\leq N_{rx}$ (n integer), the use of a weighted minimum distance detector and QAM constellations allows the Unbiased MMSE design Conventional MIMO DFE Rx to achieve the diversity vs multiplexing optimal tradeoff given by $d^*(r)$ (see [1]). $d^*(r)$ is given by the piecewise-linear function connecting the points $(k,d^*(k))$, $k=0,1,\ldots,p$, where

$$d^*(k)=(p-k)(q-k) \qquad (17)$$

with $p=\min\{N_{rx},N_{tx}\}$ and $q=\max\{N_{rx},N_{tx}\}$.
This theorem shows that the MMSE design allows to attain the optimal diversity vs. multiplexing tradeoff derived in [1].

Proof : We consider the Unbiased MMSE Conventional MIMO DFE Rx. The special symbol $\doteq$ denotes the exponential equality, *i.e.*, we write $f(\rho)\doteq\rho^b$ to denote

$$\lim_{\rho\to\infty}\frac{\ln f(\rho)}{\ln(\rho)}=b\,. \qquad (18)$$

The proof of Theorem 2 is structured in three steps. In step 1 we characterize the frame(block) error probability in term of the first symbol error probability. In step 2 we derive a lower bound on the first symbol error probability. Finally, in step 3, we characterize the behavior of the error probability for large SNR and derive the diversity versus multiplexing tradeoff. However for the lack of place in this paper we provide a shortened proof here.

Step 1:
The symbol vectors of the transmitted frame are detected sequentially using the DFE Rx. We denote by E_k the event of making an error when detecting the k^{th} symbol vector $\mathbf{b}_k$ (E_k^c is the complement or the event where no error is made when detecting the k^{th} symbol vector). Whenever there is an error on any of the detected symbols, the frame is said to be in error. P_e denotes the frame error probability.
P_e is the probability of the union of individual error events E_k, $k = 1, \ldots, T - N_{tx} + 1$,

$$P_e = P(\cup_{k=1}^{T-N_{tx}+1} E_k). \tag{19}$$

Using the following expansion

$$P_e = \sum_{k=1}^{T-N_{tx}+1} P(E_k, E_1^c, E_2^c, \ldots, E_{k-1}^c)\,, \tag{20}$$

(E_k^c is the complement event of E_k or the event where no error is made when detecting the k^{th} symbol vector) we prove that the error probability is bounded by

$$P(E_1) \leq P_e \leq (T - N_{tx} + 1)P(E_1)\,. \tag{21}$$

T is finite then $P_e \dot{\leq} P(E_1)$. The desired result then follows

$$P_e \doteq P(E_1). \tag{22}$$

Step 2:
In this step of the proof, we derive a lower bound on the first symbol error probability for a fixed channel realization $P(E_1|\mathbf{H})$.

We use the weighted minimum distance detector at the output of the unbiased MMSE Conventional MIMO DFE (16).
An error occurs if there is $\mathbf{b}_1' \neq \mathbf{b}_1$, and we decide $\mathbf{b}_1'$ for transmitted $\mathbf{b}_1$.
For an error to occur we need to have

$$\begin{aligned} ||\widehat{\mathbf{b}}_1^U - \mathbf{b}_1'||^2_{\mathbf{C}^{-1}_{\tilde{\mathbf{e}}^U\tilde{\mathbf{e}}^U}} &= ||\widehat{\mathbf{b}}_1 - (I - \tfrac{1}{\rho}\mathbf{M}^{-1})\mathbf{b}_1'||^2_{\mathbf{C}^{1}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}} \\ &\leq ||\widehat{\mathbf{b}}_1 - (I - \tfrac{1}{\rho}\mathbf{M}^{-1})\mathbf{b}_1||^2_{\mathbf{C}^{-1}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}}, \end{aligned} \tag{23}$$

where $\mathbf{C}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}} = \sigma_v^2\mathbf{M}^{-1}(I - \frac{1}{\rho}\mathbf{M}^{-1})$ and $\mathbf{M} = \mathbf{Q}^H\mathbf{D}\mathbf{Q}$.
$\mathbf{D}$ is the diagonal part of the LDU decomposition of $\mathbf{H}^H\mathbf{H} + \frac{1}{\rho}\mathbf{I}$ (verifies $\det(\mathbf{D}) = \det(\mathbf{H}^H\mathbf{H} + \frac{1}{\rho}\mathbf{I})$, see section 2.1).
We denote by $\Delta\mathbf{b} = \mathbf{b}_1 - \mathbf{b}_1'$, then (23) is equivalent to

$$\begin{aligned} &\Delta\mathbf{b}^H(\mathbf{I} - \tfrac{1}{\rho}\mathbf{M}^{-1})\mathbf{C}^{-1}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}(I - \tfrac{1}{\rho}\mathbf{M}^{-1})\Delta\mathbf{b} \\ &\leq 2\Re\{\Delta\mathbf{b}^H(\mathbf{I} - \tfrac{1}{\rho}\mathbf{M}^{-1})\mathbf{C}^{-1}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\tilde{\mathbf{e}}_1\}. \end{aligned} \tag{24}$$

Let $\Delta\mathbf{c} = \mathbf{Q}\Delta\mathbf{b}$ and $\widetilde{\mathbf{v}}_1 = \mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\tilde{\mathbf{e}}_1 = \sigma_v^{-1}\mathbf{D}^{1/2}(\mathbf{I} - \frac{1}{\rho}\mathbf{D}^{-1})^{-1/2}\mathbf{Q}\tilde{\mathbf{e}}_1$.
$\widetilde{\mathbf{v}}_1$ is spatially white: $\mathbf{C}_{\widetilde{\mathbf{v}}_1\widetilde{\mathbf{v}}_1} = \mathbf{I}_{N_{tx}}$.
Using the Cauchy-Swartz inequality [5], we show that

$$\begin{aligned} ||\widetilde{\mathbf{v}}_1||_2^2 &\geq \tfrac{1}{4\sigma_v^2}\Delta\mathbf{c}^H(\mathbf{D} - \tfrac{1}{\rho}\mathbf{I})\Delta\mathbf{c} \\ &= \tfrac{1}{4\sigma_v^2\rho}\left(\Delta\mathbf{c}^H\rho\mathbf{D}\Delta\mathbf{c} - \Delta\mathbf{b}^H\Delta\mathbf{b}\right) \end{aligned} \tag{25}$$

The channel mutual information is $I(H) = \ln\det(\mathbf{I} + \rho\mathbf{H}^H\mathbf{H}) = \ln\det(\rho\mathbf{D})$. $\mathbf{D}$ is diagonal, then using the Jensen's inequality we get that

$$\begin{aligned} \tfrac{1}{N_{tx}}\Delta\mathbf{c}^H(\rho\mathbf{D})\Delta\mathbf{c} &\geq (\textstyle\prod_{i=1}^{N_{tx}}\rho\mathbf{D}_{ii}|\Delta\mathbf{c}_i|^2)^{\frac{1}{N_{tx}}} \\ &= e^{\frac{I(H)}{N_{tx}}}(\textstyle\prod_{i=1}^{N_{tx}}|\Delta\mathbf{c}_i|^2)^{\frac{1}{N_{tx}}}. \end{aligned} \tag{26}$$

We consider a scheme where the transmitted rate varies with the SNR.
The different component of $\mathbf{b}_k$ comes from the same QAM constellation of size $(2M)^2 = \rho^{\frac{r}{N_{tx}}}$, $(r \geq 0)$ where $R(\rho) = r\ln\rho$ is the overall allocated rate and M is a positive integer. The minimum distance of the constellation is $2d$, with $d^2 = \frac{3\sigma_b^2}{2(\rho^{\frac{r}{N_{tx}}}-1)}$.
For $i = 1, \ldots, N_{tx}$: $\Delta\mathbf{b}_i = 2d(l' + jp')$, $l', p' \in \{-2M + 1, -2M + 2, \ldots, 2M - 1\}$. Then

$$\Delta\mathbf{b}^H\Delta\mathbf{b} \leq N_{tx}4d^2((2M-1)^2 + (2M-1)^2) \leq 8d^2N_{tx}\rho^{\frac{r}{N_{tx}}}. \tag{27}$$

In the other hand the choice of $\mathbf{Q}$ ensures that [2]

$$(\prod_{i=1}^{N_{tx}}|\Delta\mathbf{c}_i|^2)^{\frac{1}{N_{tx}}} \geq \frac{4d^2}{N_{tx}}. \tag{28}$$

Applying these bounds to (25), we can finally conclude that the error event for a given channel realization is included in the following event

$$\begin{aligned} ||\widetilde{\mathbf{v}}_1||_2^2 &\geq \tfrac{1}{4\sigma_v^2\rho}\left(N_{tx}(e^{\frac{I(H)}{N_{tx}}}\tfrac{4d^2}{N_{tx}}) - 8d^2N_{tx}\rho^{\frac{r}{N_{tx}}}\right) \\ &= \tfrac{d^2}{\sigma_v^2\rho}\left(e^{\frac{I(H)}{N_{tx}}} - 2N_{tx}\rho^{\frac{r}{N_{tx}}}\right) \\ &= \tfrac{3}{2(\rho^{\frac{r}{N_{tx}}}-1)}\left(e^{\frac{I(H)}{N_{tx}}} - 2N_{tx}\rho^{\frac{r}{N_{tx}}}\right) \\ &= \tfrac{3}{2(1-\rho^{-\frac{r}{N_{tx}}})}\left(e^{\frac{I(H)-r\ln\rho}{N_{tx}}} - 2N_{tx}\right) = \gamma(H) \end{aligned} \tag{29}$$

For a given channel realization, the error event E_1 is included in the event of equation (29), then

$$P(E_1|\mathbf{H}) \leq P(||\widetilde{\mathbf{v}}_1||_2^2 \geq \gamma(H)|\mathbf{H}). \tag{30}$$

$\widetilde{\mathbf{v}}_1$ can be written as

$$\begin{aligned} \widetilde{\mathbf{v}}_1 &= \mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\tilde{\mathbf{e}}_1 \\ &= \mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\mathbf{M}^{-1}\mathbf{B}^{-\dagger}(q)\mathbf{G}^{\dagger}(q)\boldsymbol{v}_1 - \mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\tfrac{1}{\rho}\mathbf{M}^{-1}(\mathbf{B}^{-\dagger}(q) - I)\mathbf{b}_1 \\ &= \widetilde{\widetilde{\mathbf{v}}}_1^1 + \widetilde{\widetilde{\mathbf{v}}}_1^2 \end{aligned} \tag{31}$$

where $\widetilde{\widetilde{\mathbf{v}}}_1^1 = \mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\mathbf{M}^{-1}\mathbf{B}^{-\dagger}(q)\mathbf{G}^{\dagger}(q)\boldsymbol{v}_1$ and
$\widetilde{\widetilde{\mathbf{v}}}_1^2 = -\mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\frac{1}{\rho}\mathbf{M}^{-1}(\mathbf{B}^{-\dagger}(q) - I)\mathbf{b}_1$. By applying one of the vector norm properties [5], we have $||\widetilde{\mathbf{v}}_1||_2 \leq ||\widetilde{\widetilde{\mathbf{v}}}_1^1||_2 + ||\widetilde{\widetilde{\mathbf{v}}}_1^2||_2$. $\widetilde{\widetilde{\mathbf{v}}}_1^2$ can be written as $\widetilde{\widetilde{\mathbf{v}}}_1^2 = -\frac{1}{\rho}\mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\mathbf{M}^{-1}\bar{\mathbf{B}}\bar{\mathbf{b}}_1$ where
$\bar{\mathbf{B}} = [\mathbf{B}_1|\mathbf{B}_2|\ldots|\mathbf{B}_{N_{tx}-1}]$ and $\bar{\mathbf{b}}_1 = [\mathbf{b}_2^T, \mathbf{b}_3^T, \ldots, \mathbf{b}_{N_{tx}}^T]^T$.
Another matrix norm property [5] is
$||\widetilde{\widetilde{\mathbf{v}}}_1^2||_2 \leq ||\frac{\sigma_b}{\rho}\mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\mathbf{M}^{-1}\bar{\mathbf{B}}||_2||\frac{1}{\sigma_b}\bar{\mathbf{b}}_1||_2$.
We have that $\frac{\sigma_b}{\rho}\mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\mathbf{M}^{-1}\bar{\mathbf{B}}(\frac{\sigma_b}{\rho}\mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\mathbf{M}^{-1}\bar{\mathbf{B}})^H = \mathrm{E}\widetilde{\widetilde{\mathbf{v}}}_1^2\widetilde{\widetilde{\mathbf{v}}}_1^{2^H} \leq \mathrm{E}\widetilde{\widetilde{\mathbf{v}}}_1^2\widetilde{\widetilde{\mathbf{v}}}_1^{2^H} + \mathrm{E}\widetilde{\widetilde{\mathbf{v}}}_1^1\widetilde{\widetilde{\mathbf{v}}}_1^{1^H} \leq \mathrm{E}\widetilde{\mathbf{v}}_1\widetilde{\mathbf{v}}_1^H \leq \mathbf{I}$.
By consequence $||\frac{\sigma_b}{\rho}\mathbf{C}^{-1/2}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\mathbf{M}^{-1}\bar{\mathbf{B}}||_2 \leq 1$.
In the other hand, all the component of $\bar{\mathbf{b}}_1$ belong to the same QAM constellation, hence
$||\frac{1}{\sigma_b}\bar{\mathbf{b}}_1||_2^2 \leq \frac{1}{\sigma_b^2}N_{tx}(N_{tx}-1)2d^2(2M-1)^2 = 3N_{tx}(N_{tx}-1)\frac{(2M-1)^2}{(2M)^2-1} \leq 3N_{tx}(N_{tx}-1) = \gamma_1$.
We conclude that $||\widetilde{\widetilde{\mathbf{v}}}_1^2||_2 \leq \gamma_1$ and $||\widetilde{\mathbf{v}}_1||_2 \leq ||\widetilde{\widetilde{\mathbf{v}}}_1^1||_2 + \gamma_1$.

Equation (30) becomes now

$$P(E_1|\mathbf{H}) \leq P(||\widetilde{\widetilde{\mathbf{v}}}_1^1||_2 \geq \sqrt{\gamma(H)} - \gamma_1|\mathbf{H}). \quad (32)$$

$\widetilde{\widetilde{\mathbf{v}}}_1^1$ has an unbiased Gaussian distribution, with covariance that is majorized by the identity

$$\mathrm{E}\widetilde{\widetilde{\mathbf{v}}}_1^1\widetilde{\widetilde{\mathbf{v}}}_1^{1\,H} \leq \mathrm{E}\widetilde{\mathbf{v}}_1^1\widetilde{\mathbf{v}}_1^{1\,H} + \mathrm{E}\widetilde{\widetilde{\mathbf{v}}}_1^2\widetilde{\widetilde{\mathbf{v}}}_1^{2\,H} = \mathrm{E}\widetilde{\mathbf{v}}_1\widetilde{\mathbf{v}}_1^H \leq \mathbf{I} \quad (33)$$

Denote $\mathbf{n}_1 = (\mathrm{E}\widetilde{\widetilde{\mathbf{v}}}_1^1\widetilde{\widetilde{\mathbf{v}}}_1^{1\,H})^{-1/2}\widetilde{\widetilde{\mathbf{v}}}_1^1$, as $(\mathrm{E}\widetilde{\widetilde{\mathbf{v}}}_1^1\widetilde{\widetilde{\mathbf{v}}}_1^{1\,H})^{1/2} \leq \mathbf{I}$ we can write the following inequality

$$||\mathbf{n}_1||_2 \geq ||\widetilde{\widetilde{\mathbf{v}}}_1^1||_2 \quad (34)$$

where $\mathbf{n}_1$ follows the unbiased Gaussian distribution with covariance identity.
The error probability is then majorized by

$$\begin{aligned} P(E_1|\mathbf{H}) &\leq P(||\widetilde{\widetilde{\mathbf{v}}}_1^1||_2 \geq \sqrt{\gamma(H)} - \gamma_1|\mathbf{H}) \\ &\leq P(||\mathbf{n}_1||_2 \geq \sqrt{\gamma} - \gamma_1|\mathbf{H}) \\ &= P(||\mathbf{n}_1||_2^2 \geq \gamma_2(H)|\mathbf{H}), \end{aligned} \quad (35)$$

where $\gamma_2(H) = (\sqrt{\gamma(H)} - \gamma_1)^2$.
Step 3:
In step 3 we seek to study the behavior of the error probability for large SNR, in order to derive the diversity versus multiplexing tradeoff achieved by our scheme.
We define $p = \min\{N_{rx}, N_{tx}\}$, $q = \max\{N_{rx}, N_{tx}\}$ and λ_i, $i = 1, \ldots, p$, to be the nonzero eigenvalues of $\mathbf{H}^H\mathbf{H}$ sorted in the increasing order.
We continue in the foot steps of [1] and use the following variable change $\lambda_i \triangleq \rho^{-\alpha_i}$. At high SNR we have $(1 + \rho\lambda_i) \doteq \rho^{(1-\alpha_i)^+}$, where $(x)^+$ denotes $\max\{0, x\}$. In the other hand the mutual information verifies $I(H) = \sum_{i=1}^p \ln(1 + \rho\lambda_i)$, hence $e^{I(H)} \doteq \rho^{\sum_{i=1}^p (1-\alpha_i)^+}$.

In [1], it was shown that for an allocated rate $r \ln \rho$, the outage probability is

$$P(outage) \doteq P(\sum_{i=1}^p (1 - \alpha_i)^+ \leq r) \doteq \rho^{-d_{out}(r)}, \quad (36)$$

where $d_{out}(r)$ is given by the piecewise-linear function connecting the points $(k, d_{out}(k))$, $k = 0, 1, \ldots, p$, where

$$d_{out}(k) = (p - k)(q - k)\,. \quad (37)$$

It was also shown in the same paper that any scheme with rate $R(\rho) = r \ln \rho$ has an error probability that verifies

$$P_e \dot{\geq} \rho^{-d_{out}(r)}, \quad (38)$$

$d_{out}(r) = d^*(r)$ is also called the optimal tradeoff curve.

Let ϵ be a small real positive number $\epsilon > 0$. We define the $outage_\epsilon$ event for $\sum_{i=1}^{N_{tx}} (1 - \alpha_i)^+ \leq r + \epsilon$. The compliment event of $outage_\epsilon$ is denoted as $no\,outage_\epsilon$.
Then the following relation is verified

$$\{outage_\epsilon\} \cup E_1 = \{outage_\epsilon\} \cup (\{no\,outage_\epsilon\} \cap E_1). \quad (39)$$

A upper bound on $P(E_1)$ can then be derived as

$$\begin{aligned} P(E_1) &\leq P(\{outage_\epsilon\} \cup E_1) \\ &= P(outage_\epsilon) + P(E_1, no\,outage_\epsilon). \end{aligned} \quad (40)$$

For (36) we conclude that

$$P(outage_\epsilon) \doteq P(\sum_{i=1}^p (1 - \alpha_i)^+ \leq r + \epsilon) \doteq \rho^{-d_{out}(r+\epsilon)}, \quad (41)$$

We want to characterize $P(E_1, no\,outage_\epsilon)$. By applying the Chernoff bound to (35), and for any $\lambda > 0$, we get

$$\begin{aligned} P(E_1,\, no\,outage_\epsilon) &= \int_{no\,outage_\epsilon} P(E_1|\mathbf{H})\, f(\mathbf{H})\, d\mathbf{H} \\ &\leq \int_{no\,outage_\epsilon} P(||\mathbf{n}_1||_2^2 \geq \gamma_2(H)|\mathbf{H})\, f(\mathbf{H})\, d\mathbf{H} \\ &\leq \int_{no\,outage_\epsilon} (1 - \lambda)^{-N_{tx}} e^{-\lambda\gamma_2(H)}\, f(\mathbf{H})\, d\mathbf{H} \\ \text{for } \lambda = \tfrac{1}{2} \quad &\leq \int_{no\,outage_\epsilon} 2^{N_{tx}} e^{-\frac{\gamma_2(H)}{2}}\, f(\mathbf{H})\, d\mathbf{H}\,, \end{aligned} \quad (42)$$

For any realization of the channel with $no\,outage_\epsilon$ verifies $\sum_{i=1}^{N_{tx}} (1 - \alpha_i)^+ > r + \epsilon$, then

$$P(E_1,\, no\,outage_\epsilon) \dot{\leq} \int_{no\,outage_\epsilon} \gamma_\epsilon(\rho)\;\; f(\mathbf{H})\, d\mathbf{H} \;\leq \gamma_\epsilon(\rho)\,, \quad (43)$$

where $\gamma_\epsilon(\rho) = 2^{N_{tx}} e^{-\left(\left(\frac{3}{2(1-\rho^{-\frac{r}{N_{tx}}})}(\rho^{\frac{\epsilon}{N_{tx}}} - 2N_{tx})\right)^{\frac{1}{2}} - \gamma_1 \right)^2 /2}$.
For any $\epsilon > 0$ the following property is verified $\lim_{\rho\to\infty} \frac{\ln \gamma_\epsilon(\rho)}{\ln \rho} = -\infty$. Hence for any finite y we have $P(E_1,\, no\,outage_\epsilon) \dot{\leq} \rho^{-y}$, and by consequence

$$P(E_1,\, no\,outage_\epsilon) \dot{\leq} \rho^{-d_{out}(r+\epsilon)}. \quad (44)$$

Combining this result with (40) and (41) leads to

$$P(E_1) \dot{\leq} \rho^{-d_{out}(r+\epsilon)}, \quad (45)$$

which is valid for any $\epsilon > 0$, we have then

$$P(E_1) \dot{\leq} \rho^{-d_{out}(r)} = \rho^{-d^*(r)}. \quad (46)$$

Using (22), we end with an upper bound to the frame error probability

$$P_e \dot{\leq} \rho^{-d^*(r)}. \quad (47)$$

The lower bound of (38), allows us to finally conclude that our scheme attain the optimal diversity vs multiplexing tradeoff, or

$$P_e \doteq \rho^{-d^*(r)} \quad (48)$$

□

4. REFERENCES

[1] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Info. Theory*, vol. 49, no. 5, pp. 1073 –1096, May 2003.

[2] A. Medles and D.T.M. Slock, "Linear Convolutive Space-Time Precoding for Spatial Multiplexing MIMO Systems," in *Proc. 39th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, Illinois, Oct. 2001.

[3] A. Medles and D.T.M. Slock, "Miltistream Space-Time Coding by Spatial Spreading, Scrambling and Delay Diversity," in *Proc. ICASSP Conf.*, Orlando, FL, May 2002.

[4] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 2000.

[5] G.H. Golub and C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, third edition edition, 1996.