OPTIMAL DIVERSITY VS MULTIPLEXING TRADEOFF FOR FREQUENCY SELECTIVE MIMO CHANNELS

Abdelkader Medles

Bell Labs, Lucent Technologies
The Quadrant, Stonehill Green
Swindon, SN5 7DJ, UK
Email: medles@lucent.com

Dirk T.M. Slock

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

ABSTRACT

In this paper we derive the optimal diversity versus multiplexing tradeoff for a frequency selective i.i.d. Rayleigh MIMO channel. This tradeoff is shown to be better than the one of the frequency flat MIMO channel. Traditional approaches for frequency selective channels use OFDM techniques in order to exploit the diversity gain due to frequency selectivity. We show that although coding in OFDM over a sufficient subset of subcarriers allows to exploit full diversity as such (at fixed rate), such an approach leads to a suboptimal diversity vs multiplexing tradeoff.

1. INTRODUCTION

Frequency selective channels, also known as inter-symbol interference (ISI) channels, are regarded in general as troublesome channels from a performance point of view and practical systems assume the use of the OFDM approach to avoid this type of channels. However, recent studies [1, 2] show that SISO/SIMO frequency selective channels improve performance and allow to achieve better diversity versus multiplexing gain compared to frequency-flat channels. In this paper, we generalize the optimal diversity versus multiplexing tradeoff (introduced first for flat MIMO channels in [3]) to frequency selective MIMO channels. The tradeoff obtained provides at any rate more diversity gain than in the case of flat channels. Furthermore, this tradeoff is better than the one obtained via coding over a set of linearly independent or even uncorrelated subcarriers of an OFDM system. Such a result is very surprising and suggests to stay in the time domain or code across more subcarriers for more efficient diversity exploitation.

Consider linear digital modulation over a linear channel with additive white circular complex Gaussian noise with variance σ_v^2. The number of antennas is N_t at the transmitter side and N_r at the receiver side. The channel is frequency selective, with a delay spread L. The complex received signal at antenna r is

$$\mathbf{y}_k^r = \sum_{l=0}^{L-1}\sum_{t=1}^{N_t} \mathbf{H}_l^{rt}\mathbf{x}_{k-l}^t + \mathbf{v}_k^r \tag{1}$$

where $\mathbf{x}_k^t$ is the transmitted signal at antenna t at time k, $\mathbf{H}_l^{rt},\ l = 0,\ldots,\ L-1$ are the complex coefficients of the channel impulse response from antenna t to antenna r, and $\mathbf{v}_k^r$ is the noise at antenna r. Collecting the received signals in one vector symbol $\mathbf{y}_k = [\mathbf{y}_k^1,\ldots,\mathbf{y}_k^{N_r}]^T$, we can then write

$$\mathbf{y}_k = \sum_{l=0}^{L-1} \mathbf{H}_l\mathbf{x}_{k-l} + \mathbf{v}_k \tag{2}$$

where $\mathbf{x}_k = [\mathbf{x}_k^1,\ldots,\mathbf{x}_k^{N_t}]^T$, $\mathbf{H}_l : N_r \times N_t,\ l = 0,\ldots,\ L-1$ are the matricial coefficients of the channel impulse response, and $\mathbf{v}_k = [\mathbf{v}_k^1,\ldots,\mathbf{v}_k^{N_r}]^T$, $\mathbf{v}_k \sim \mathcal{CN}(0,\sigma_v^2\mathbf{I}_{N_r})$. Superscripts ^{T}, ^{H} denote transpose and Hermitian transpose respectively.

1.1. Block (Frame) Transmission

We consider a transmission over a large block of length $T >> L$, the channel is assumed to be constant over the block, and varying independently from one block to another. The received signal can be written as

$$\mathbf{Y} = \mathcal{A}_T(\mathbf{H})\mathbf{X} + \mathbf{V}\,, \tag{3}$$

where $\mathbf{Y} = [\mathbf{y}_1^T,\mathbf{y}_2^T,\ldots,\mathbf{y}_{T+L-1}^T]^T$, $\mathbf{X} = [\mathbf{x}_1^T,\mathbf{x}_2^T,\ldots,\mathbf{x}_T^T]^T$ and $\mathbf{V} = [\mathbf{v}_1^T,\mathbf{v}_2^T,\ldots,\mathbf{v}_{T+L-1}^T]^T$. $\mathcal{A}_T(\mathbf{H})$ is the following $N_r(T+L-1)\times N_tT$ block Toeplitz matrix

$$\mathcal{A}_T(\mathbf{H}) = \begin{bmatrix} \mathbf{H}_0 & 0 & \ldots & 0 \\ \mathbf{H}_1 & \mathbf{H}_0 & \ddots & \vdots \\ \vdots & \mathbf{H}_1 & \ddots & 0 \\ \mathbf{H}_{L-1} & \vdots & \ddots & \mathbf{H}_0 \\ 0 & \mathbf{H}_{L-1} & \vdots & \mathbf{H}_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \mathbf{H}_{L-1} \end{bmatrix}. \tag{4}$$

For a general input signal the mutual information per symbol period, is

$$I_T(\mathbf{H}) = \frac{1}{T+L-1} I(\mathbf{X};\mathbf{Y}|\mathbf{H})\,. \tag{5}$$

Let $\mathbf{R_{XX}} = \mathrm{E}\ \mathbf{XX}^H$ be the input correlation. The power constraint is given by $\mathrm{tr}\,\mathbf{R_{XX}} \le T\ P$, where P is the maximum transmit power per symbol period. For a given color of the transmitted signal ($\mathbf{R_{XX}}$), the mutual information per symbol period is maximized for a Gaussian centered input, for which it is given by

$$I_T(\mathbf{H}) = \frac{1}{T+L-1}\ln\det(\mathbf{I}_{N_r(T+L-1)} + \frac{1}{\sigma_v^2}\mathcal{A}_T(\mathbf{H})\mathbf{R_{XX}}\mathcal{A}_T^H(\mathbf{H}))\,. \tag{6}$$

Stationary Input

For large block length $T >> L$, we consider the case of stationary input. We can then define the second order statistics of the input

The Eurecom Institute'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.

$\mathbf{S}_i = \mathrm{E}\,\mathbf{x}_k\mathbf{x}_{k-i}^H$. The power spectral density function of the input is then $\mathbf{S}(z) = \sum_i \mathbf{S}_i z^{-i}$ ($\mathbf{S}(z)$ is the z-transform of $\mathbf{S}_i$). $\mathbf{R_{XX}}$ is now a block Toeplitz matrix of $\mathbf{S}_i$:

$$\mathbf{R_{XX}} = \begin{bmatrix} \mathbf{S}_0 & \mathbf{S}_1 & \dots & \mathbf{S}_{L-2} & \mathbf{S}_{L-1} \\ \mathbf{S}_1^H & \mathbf{S}_0 & \ddots & \ddots & \mathbf{S}_{L-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{S}_{L-2}^H & \ddots & \ddots & \mathbf{S}_0 & \mathbf{S}_1 \\ \mathbf{S}_{L-1}^H & \mathbf{S}_{L-2}^H & \dots & \mathbf{S}_1^H & \mathbf{S}_0 \end{bmatrix}. \quad (7)$$

Frequency Selective Block Fading Rayleigh Model
We construct this model as a generalization of the flat Rayleigh MIMO channel model. $\mathbf{H}_l,\ l = 0, \dots, L-1$ are i.i.d, and each $\mathbf{H}_l$ has i.i.d. Gaussian components $\mathbf{H}_l^{rt} \sim \mathcal{CN}(0, \frac{1}{L})$ for $1 \le r \le N_r$ and $1 \le t \le N_t$.

1.2. Continuous transmission ($T \to \infty$)

The mutual information $I(\mathbf{H})$ in this case corresponds to the limit of $I_T(\mathbf{H})$, and is given in the following lemma.
Lemma 1: The limit of $I_T(\mathbf{H})$, $I(\mathbf{H}) = \lim_{T\to+\infty} I_T(\mathbf{H})$, is

$$I(\mathbf{H}) = \frac{1}{2\pi j}\oint \frac{dz}{z} \ln\det(\mathbf{I} + \frac{1}{\sigma_v^2}\mathbf{H}(z)\mathbf{S}(z)\mathbf{H}^\dagger(z)). \quad (8)$$

Proof: This is a generalization of the SISO case shown in [4], the proof is similar.

1.3. Bounds on the mutual information (MI) for white input

Let $J_T(\mathbf{H})$ denote the MI for spatiotemporally white input,

$$J_T(\mathbf{H}) = \frac{1}{T+L-1}\ln\det(\mathbf{I}_{N_r(T+L-1)} + \rho\,\mathcal{A}_T(\mathbf{H})\mathcal{A}_T^H(\mathbf{H}))\,. \quad (9)$$

where $\rho = \frac{\sigma_x^2}{\sigma_v^2} = \frac{P}{N_t\sigma_v^2}$. For large block length T, $J_T(\mathbf{H})$ converges to

$$J(\mathbf{H}) = \frac{1}{2\pi j}\oint \frac{dz}{z}\ln\det(\mathbf{I}_{N_r} + \rho\,\mathbf{H}(z)\mathbf{H}^\dagger(z))\,. \quad (10)$$

For finite block length, the following lemma gives useful bounds on $J_T(\mathbf{H})$ in terms of the infinite block length mutual information ($J(\mathbf{H})$).
Lemma 2 : For white input, the mutual information of a block of length T is bounded by

$$(1 - \tfrac{L-1}{T+L-1})J(\mathbf{H}) \le J_T(\mathbf{H}) \le J(\mathbf{H})\,. \quad (11)$$

Proof : The proof is omitted for lack of space.

2. DIVERSITY AND MULTIPLEXING FOR FREQUENCY SELECTIVE MIMO CHANNELS

In [3], Zheng and Tse give a new definition of the diversity and spatial multiplexing that considers adaptive SNR schemes. For the generalization of the tradeoff to frequency selective MIMO channel we consider the same definitions. Scheme $\mathcal{C}(\rho)$ is a family of codes of block length T (one for each SNR level), that supports a bit rate $R(\rho)$. This scheme is to achieve *spatial multiplexing* r and diversity gain d if the data rate satisfies

$$\lim_{\rho\to\infty}\frac{R(\rho)}{\ln(\rho)} = r\,, \quad (12)$$

and the average error probability satisfies

$$\lim_{\rho\to\infty}\frac{\ln P_e(\rho)}{\ln(\rho)} = -d\,. \quad (13)$$

For each r, $d^*(r)$ is defined to be the supremum of the diversity advantage achieved over schemes. The maximal diversity gain is defined by $d^*_{max} = d^*(0)$ and the maximal spatial multiplexing gain is $r^*_{max} = sup\{r : d^*(r) > 0\}$.
We will use the special symbol $\doteq$ to denote the exponential equality, *i.e.*, we write $f(\rho) \doteq \rho^b$ to denote

$$\lim_{\rho\to\infty}\frac{\ln f(\rho)}{\ln(\rho)} = b\,. \quad (14)$$

Zheng and Tse considered a **flat Rayleigh MIMO channel**, and using properties of the distribution of eigenvalues they showed the following results.

Optimal Tradeoff Curve for a flat MIMO channel : Assume $T \ge N_r + N_t - 1$. The optimal tradeoff curve $d^*(r)$ is given by the piecewise-linear function connecting the points $(k, d^*(k))$, $k = 0, 1, \dots, p$, where

$$d^*(k) = (p-k)(q-k)\,. \quad (15)$$

with $p = \min\{N_r, N_t\}$, $q = \max\{N_r, N_t\}$. In particular $d^*_{max} = N_tN_r$ and $r^*_{max} = \min(N_r, N_t)$.

2.1. Behavior of the diversity vs multiplexing optimal tradeoff for a frequency selective Rayleigh MIMO channel

In the case of a frequency selective channel the optimal tradeoff curve $d^*(r)$ has a behavior characterized by the following theorem.

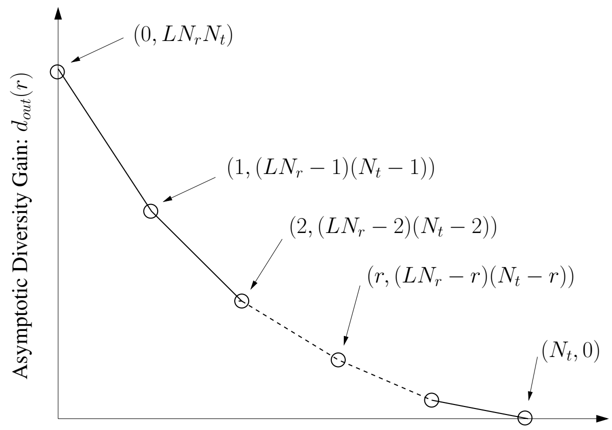

Fig. 1. Asymptotic diversity vs. multiplexing tradeoff for frequency selective channel and $N_t \le N_r$

Theorem 1: Assume $T \ge Lq + p - 1$. The optimal tradeoff curve is characterized by a diversity advantage function $d^*(r)$ which is bounded as follows:

$$d_{out}((1 + \frac{L-1}{T})r) \le d^*(r) \le d_{out}(r)\,. \quad (16)$$

$d_{out}(r)$ is the asymptotic (in T) diversity advantage given by the piecewise-linear function connecting the points $(k, d_{out}(k))$, $k =$

$0, 1, \ldots, p$, and

$$d_{out}(k) = (p-k)(Lq-k)\,. \quad (17)$$

Independently of the block length the maximum achievable diversity advantage is $d^*_{max} = d^*(0) = LN_rN_t$. However, the maximal spatial multiplexing gain r^*_{max} depends on the block length T, and is bounded by $(1 - \frac{L-1}{T+L-1})p \leq r^*_{max} \leq p$. Then asymptotically $r^*_{max} = p$.

Remarks :

- This theorem defines a region in the plane of diversity vs multiplexing, where $d^*(r)$ is situated. However, for large block length $T >> L$, this region is restricted to the curve $d_{out}(r)$, $d^*(r) = d_{out}(r)$.
- For $T >> L$ and $r = k$, $k = 0, 1, \ldots, p$, $(p-k)(Lq-k) - (p-k)(q-k) = (p-k)(L-1)q$ is the supplementary diversity gain provided by a frequency selective channel w.r.t. to a flat one with the same number of antennas.
- OFDM systems assume in general coding of the transmitted signal over L independent subcarriers. This is equivalent to coding over L independent blocks of a MIMO flat block fading channel. The optimal diversity gain is then [3] the piecewise-linear function connecting the points $(k, L(p-k)(q-k))$, $k = 0, 1, \ldots, p$. The difference $(p-k)(Lq-k) - L(p-k)(q-k) = (p-k)(L-1)k$, is the supplementary diversity gain provided by the frequency selective channel, and can be attained by exploiting the diversity in the time domain (TDMA) or by coding across more subcarriers. The exact number of needed subcarriers is left here as an open problem . Note that coding across L independent subcarriers in an OFDM system allows to attain the unreduced $d^*_{max} = LN_tN_r$, which motivated the introduction of coding over such a reduced set of subcarriers.

In the following we present the proof of Theorem 1.

2.2. Preliminary Results

Bounds on $I(\mathbf{H})$ for general input color

For a general input spectrum, under the power constraint $(\frac{1}{2\pi j}\oint \frac{dz}{z}\mathrm{tr}\{\mathbf{S}(z)\} \leq P)$, the mutual information $I(\mathbf{H})$ for infinite block length is given in (8). The following lemma gives an upper bound on $I(\mathbf{H})$.

Lemma 3: The mutual information for infinite block length is bounded as follows

$$I(\mathbf{H}) \leq \ln\det(\mathbf{I} + \rho\, LN_t\, \bar{\mathbf{H}}\bar{\mathbf{H}}^H)\,, \quad (18)$$

where $\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_0 \\ \vdots \\ \mathbf{H}_{L-1} \end{bmatrix} = [\mathbf{H}_0^T, \ldots, \mathbf{H}_{L-1}^T]^T$ in the case $N_t \leq N_r$, and $\bar{\mathbf{H}} = [\mathbf{H}_0, \mathbf{H}_1, \ldots, \mathbf{H}_{L-1}]$ for $N_t \geq N_r$ (for both cases $\bar{\mathbf{H}}$ has rank $p = \min\{N_r, N_t\}$).

Proof : The proof is omitted for lack of space.

We introduce now a paraunitary precoder for white input ($\mathbf{S}(z) = \sigma_x^2\mathbf{I}_{N_t}$) that will allow us to characterize $J(\mathbf{H})$ in (10).

Introduction of the STS precoder:

The STS precoder $\mathbf{T}(z)$ was introduced in [5]. $\mathbf{T}(z)$ is a $N_t \times N_t$ MIMO filter

$$\begin{aligned} \mathbf{T}(z) &= \mathbf{D}(z)\,\mathbf{Q} \\ \mathbf{D}(z) &= \mathrm{diag}\,\{1,\, z^{-L}, \ldots, z^{-L(N_t-1)}\}\,, \end{aligned} \quad (19)$$

where $\mathbf{Q}$ is a unitary matrix ($\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$)

$$\mathbf{Q}^s = \frac{1}{\sqrt{N_t}} \begin{bmatrix} 1 & \theta_1 & \ldots & {\theta_1}^{N_t-1} \\ 1 & \theta_2 & \ldots & {\theta_2}^{N_t-1} \\ \vdots & \vdots & & \vdots \\ 1 & \theta_{N_t} & \ldots & {\theta_{N_t}}^{N_t-1} \end{bmatrix}, \quad (20)$$

where the θ_i are the roots of $\theta^{N_t} - j = 0$, $j = \sqrt{-1}$. $\mathbf{T}(z)$ is paraunitary ($\mathbf{T}(z)\mathbf{T}^\dagger(z) = \mathbf{I}_{N_t}$), and so is $\mathbf{T}(z^L)$, hence using it as a precoder does not modify the mutual information for white Gaussian input

$$\begin{aligned} J(\mathbf{H}) &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z}\ln\det(\mathbf{I}_{N_{tx}} + \rho\,\mathbf{H}^\dagger(z)\mathbf{H}(z)) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z}\ln\det(\mathbf{I}_{N_{tx}} + \rho\,\mathbf{T}^\dagger(z)\mathbf{H}^\dagger(z)\mathbf{H}(z)\mathbf{T}(z)) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z}\ln\det(\rho\ \mathbf{R}(z)) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z}\ln\det(\rho\,\mathbf{L}^\dagger(z)\Sigma\mathbf{L}(z)) \\ &= \ln\det(\rho\,\Sigma) = \sum_{n=1}^{N_{tx}} J_n\,, \end{aligned} \quad (21)$$

where $J_n = \ln\Sigma_{nn}$ and in the fourth equality we replaced $\mathbf{R}(z) = \frac{1}{\rho}\mathbf{I}_{N_{tx}} + \mathbf{T}^\dagger(z)\mathbf{H}^\dagger(z)\mathbf{H}(z)\mathbf{T}(z)$ by its UDL matrix spectral factorization. $\mathbf{L}(z)$ is lower triangular, with diagonal elements, $\mathbf{L}_{ii}(z)$, $i = 1, \ldots, N_t$, causal, monic[1] and minimum phase[2]. The lower triangular elements $\mathbf{L}_{ij}(z)$, $i > j$, are arbitrary (non causal) transfer functions. W.r.t. a classical causal MIMO spectral factor, the degrees of freedom of the strictly causal upper triangular elements have been transferred to the anticausal part of the lower triangular elements. In the fifth equality we exploit the fact that $\det(\mathbf{L}(z)) = \prod_{n=1}^{N_t}\mathbf{L}_{nn}(z)$ is causal, monic and minimum phase. Let us denote $\mathbf{V}_n = [\mathbf{e}_n, \mathbf{e}_{n+1}, \ldots, \mathbf{e}_{N_t}]$, where $\mathbf{e}_n$ is the $N_t \times 1$ vector with 1 at the n^{th} position and all other entries are zeros. By identification from the UDL spectral factorization of $\mathbf{R}(z)$, *i.e.* $\mathbf{L}(z)^\dagger\Sigma\mathbf{L}(z) = \mathbf{R}(z)$, we show that $([\mathbf{V}_n^H\mathbf{R}(z)\mathbf{V}_n]^{-1})_{11} = \frac{1}{\Sigma_{nn}\mathbf{L}_{nn}(z)\mathbf{L}^\dagger_{nn}(z)}$, where the diagonal coefficient $\mathbf{L}_{nn}(z)$ is causal, monic and minimum phase, hence it satisfies $\frac{1}{2\pi j}\oint\frac{dz}{z}\ln(\mathbf{L}_{nn}(z)\mathbf{L}^\dagger_{nn}(z)) = 0$[3]. Finally

$$J_n = \ln\rho\Sigma_{nn} = -\frac{1}{2\pi j}\oint\frac{dz}{z}\ln([\mathbf{V}_n^H\rho\mathbf{R}(z)\mathbf{V}_n]^{-1})_{11}\,. \quad (22)$$

Bounds on $J(\mathbf{H})$:

SIMO case

Lemma 4:

For this case the mutual information $J(\mathbf{H})$ is bounded by

$$\ln(\frac{1+\rho||\underline{\mathbf{h}}||^2}{\gamma_L}) \leq \frac{1}{2\pi j}\oint\frac{dz}{z}\ln(1+\rho\mathbf{h}^\dagger(z)\mathbf{h}(z)) \leq \ln\left(1+\rho||\underline{\mathbf{h}}||^2\right) \quad (23)$$

where $\underline{\mathbf{h}} = \begin{bmatrix} \mathbf{h}_0 \\ \vdots \\ \mathbf{h}_{L-1} \end{bmatrix}$ represents the vectorized channel impulse response and $\gamma_L = \sum_{l=0}^{L-1}\binom{L-1}{l}^2$.

[1] A monic SISO filter has first coefficient equal to 1

[2] A rational $f(z)$ is said to be minimum phase if all of its poles and zeros are inside the unit circle

[3] $\frac{1}{2\pi j}\oint\frac{dz}{z}\ln(f(z)f(z)^\dagger) = 0$ for $f(z)$ causal, monic and minimum phase

Proof : The proof is omitted for lack of space.

MIMO case
Let us introduce first

$$c_n = \begin{cases} (q-n)\ln\left(\frac{\gamma_{qL}(q-n+1)}{q-n}\right) & ,1 \le n \le q-1 \\ 0 & ,n=q \end{cases} \quad (24)$$

Lemma 5:
For $N_t \ge N_r$, J_n is bounded by

$$-\ln(N_t L \gamma_{N_t L}) - c_n \le J_n - \ln(1+\rho L\, s_{N_t(L-1)+n}) \le c_n\,, \quad (25)$$

where s_n, $n = 1, \ldots, N_t L$, are the eigenvalues of $\bar{\mathbf{H}}^H \bar{\mathbf{H}}$ sorted in a nondecreasing order.
Proof : The proof is omitted for lack of space.
Lemma 6:
Using the results of lemma 5, we conclude for any N_t, N_r,

$$-q\ln(qL\gamma_{qL}) - c \le J(\mathbf{H}) - \ln\det(\mathbf{I} + \rho\, L\, \bar{\mathbf{H}}\bar{\mathbf{H}}^H) \le c \quad (26)$$

where $c = \sum_{n=1}^{q} c_n$.
Proof : For $N_t \ge N_r$, the derivation is straightforward from Lemma 5, in fact for $N_t \ge N_r$ there are only $p = N_r$ non-zero eigenvalues in $\bar{\mathbf{H}}^H\bar{\mathbf{H}}$: $s_{N_t L - N_r + 1}, \ldots, s_{N_t L}$. This ensures that $\sum_{n=1}^{N_t} \ln(1+\rho L s_{N_t(L-1)+n}) = \ln\det(\mathbf{I} + \rho\, L\, \bar{\mathbf{H}}\bar{\mathbf{H}}^H)$.
For the case $N_t \le N_r$ we observe that the instantaneous mutual information satisfies

$$\begin{aligned} J(\mathbf{H}) &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z} \ln\det(\mathbf{I} + \rho\,\mathbf{H}(z)\mathbf{H}^\dagger(z)) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z} \ln\det(\mathbf{I} + \rho\,\mathbf{H}(z)\mathbf{H}^H(z)) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z} \ln\det((\mathbf{I} + \rho\,\mathbf{H}(z)\mathbf{H}^H(z))^T) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z} \ln\det(\mathbf{I} + \rho\,\mathbf{H}^T(z)\mathbf{H}^*(z))\,. \end{aligned} \quad (27)$$

In the second equality we replaced † by ^{H} because the integral is done on the unit circle, the two operators are then equivalent. We conclude that the capacity is the same as for a new virtual channel $\mathbf{H}^T(z)$. As a consequence the result of the first case, $N_t \ge N_r$, holds also here after interchanging N_r and N_t, and replacing $\mathbf{H}_i$ by $\mathbf{H}_i^T$ for $i = 0, \ldots, L-1$.

2.3. Outage Probability Behavior

An outage event occurs when the mutual information of the channel does not support a target data rate

$$\{\mathbf{H} : I_T(\mathbf{H}) = \frac{1}{T+L-1} I(\mathbf{X};\mathbf{Y}|\mathbf{H}) < R\}\,. \quad (28)$$

The mutual information depends on the input distribution and the channel realization. Since the mutual information is maximized for Gaussian inputs, we can then define the outage probability as

$$P_{out}(R) = \min_{\mathbf{R_{XX}} \ge 0,\, \mathrm{tr}\{\mathbf{R_{XX}}\} \le T\,P} P(I_T(\mathbf{H}) < R). \quad (29)$$

Any covariance matrix $\mathbf{R_{XX}}$ in the set $\{\mathbf{R_{XX}} \ge 0,\ \mathrm{tr}\{\mathbf{R_{XX}}\} \le T\,P\}$ satisfies $\mathbf{R_{XX}} \le T\,P\,\mathbf{I}_{TN_t}$, where $\mathbf{I}_{TN_t}$ is the identity matrix of size $TN_t \times TN_t$.
From (6) we can then write that

$$I_T(\mathbf{H}) \le \frac{1}{T+L-1} \ln\det(\mathbf{I} + N_t T \rho \mathcal{A}_T(\mathbf{H})\mathcal{A}_T^H(\mathbf{H}))\,. \quad (30)$$

The outage probability then satisfies

$$\begin{aligned} &P[\tfrac{1}{T+L-1}\ln\det(\mathbf{I} + N_t T\rho\mathcal{A}_T(\mathbf{H})\mathcal{A}_T^H(\mathbf{H})) < R] \\ &\le P_{out}(R) \le P[\tfrac{1}{T+L-1}\ln\det(\mathbf{I} + \rho\mathcal{A}_T(\mathbf{H})\mathcal{A}_T^H(\mathbf{H})) < R]\,, \end{aligned} \quad (31)$$

where for the upper bound we have just chosen $\mathbf{R_{XX}} = \frac{P}{N_t}\mathbf{I}_{TN_t}$. As has been observed in [3], the two constant scalings ($N_t T$ and 1) of ρ lead to the same exponential equality, hence

$$P_{out}(R) \doteq P[\frac{1}{T+L-1}\ln\det(\mathbf{I} + \rho\mathcal{A}_T(\mathbf{H})(\mathcal{A}_T(\mathbf{H}))^H) < R]\,. \quad (32)$$

Now, using the results of Lemma 2, namely (11), we get

$$P[J(\mathbf{H}) < R] \dot{\le} P_{out}(R) \dot{\le} P[(1 - \frac{L-1}{T+L-1})J(\mathbf{H}) < R]\,. \quad (33)$$

Finally, by the exploitation of Lemma 6, or hence (26), we conclude

$$\begin{aligned} &P[\ln\det(\mathbf{I} + \rho\,\bar{\mathbf{H}}\bar{\mathbf{H}}^H) < R] \dot{\le} P_{out}(R) \dot{\le} \\ &P[(1 - \tfrac{L-1}{T+L-1})\ln\det(\mathbf{I} + \rho\,\bar{\mathbf{H}}\bar{\mathbf{H}}^H) < R]\,. \end{aligned} \quad (34)$$

Remarks :

- The scale factor $N_t T$ in (30) is very loose for large block length ($T >> L$), and no more valid for $T \to \infty$. However, for such cases practical signals are stationary and admit a spectrum $\mathbf{S}(z)$, $I_T(\mathbf{H})$ then converges to the mutual information of infinite block length $I(\mathbf{H})$ (8). We can then use the upper bound of (18) which leads directly to the lower bound in (34). By consequence the reasoning is still valid for infinite block length T, with the upper bound in (34) remaining the same.
- For large block length ($T >> L$), (34) becomes $P_{out}(R) \doteq P[\ln\det(\mathbf{I} + \rho\,\bar{\mathbf{H}}\bar{\mathbf{H}}^H) < R]$. The frequency selective channel behaves as the flat MIMO channel $\bar{\mathbf{H}} = [\mathbf{H}_0^T, \ldots, \mathbf{H}_{L-1}^T]^T$ in the case $N_t \le N_r$ and $\bar{\mathbf{H}} = [\mathbf{H}_0, \mathbf{H}_1, \ldots, \mathbf{H}_{L-1}]$ for $N_t \ge N_r$.

Let the data rate $R = r\ln\rho$. In [3] it was shown that for a $N_r \times N_t$ flat Rayleigh distributed MIMO channel

$$P[\ln\det(\mathbf{I} + \rho\,\mathbf{H}\mathbf{H}^H) < r\ln\rho] \doteq \rho^{-d_{out}(r)}\,, \quad (35)$$

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

$$d_{out}(k) = (p-k)(q-k)\,. \quad (36)$$

Recall that $p = \min\{N_r, N_t\}$, $q = \max\{N_r, N_t\}$.
In our case, $\bar{\mathbf{H}}$ has the same probability distribution as that of a flat Rayleigh distributed MIMO channel of size $N_r L \times N_t$ for $N_t \le N_r$, and $N_r \times N_t L$ for $N_t \ge N_r$. These two cases lead to the same result

$$P[\ln\det(\mathbf{I} + \rho\,\bar{\mathbf{H}}\bar{\mathbf{H}}^H) < r\ln\rho] \doteq \rho^{-d_{out}(r)}\,, \quad (37)$$

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

$$d_{out}(k) = (p-k)(Lq-k)\,. \quad (38)$$

Using this results it is straightforward to derive the following theorem.

Theorem 2: For a frequency selective and i.i.d. Rayleigh distributed MIMO channel, the outage probability satisfies

$$\rho^{-d_{out}(r)} \dot{\leq} P_{out}(R) \dot{\leq} \rho^{-d_{out}((1+\frac{L-1}{T})r)}\,, \qquad (39)$$

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

$$d_{out}(k) = (p-k)(Lq-k)\,. \qquad (40)$$

Remark : From (39) we can see that for large block $T >> L$, the outage probability satisfies $P_{out}(R) \doteq \rho^{-d_{out}(r)}$.

2.4. Proof of Theorem 1

In order to prove Theorem 1, we characterize the error probability of any coding scheme (for a given rate $R = r\ln\rho$) by $P_e(\rho) = \rho^{-d(r)}$. We recall that $d^*(r)$ is the supremum of the diversity advantage $d(r)$ achieved over all these schemes.
To show the upper bound in (16), we can use the same proof of lemma 5 in [3]. This will lead to the following result

$$P_e(\rho) \dot{\geq} \rho^{-d_{out}(r)}\,. \qquad (41)$$

Or in other terms, for any scheme $d(r) \leq d_{out}(r)$. Hence, the supremum verifies

$$d^*(r) \leq d_{out}(r)\,. \qquad (42)$$

Proof of the lower bound in Theorem 1 : We need now to prove that $d_{out}((1+\frac{L-1}{T})r) \leq d^*(r)$, for $T \geq Lq+p-1$. To show this result we provide an upper bound on the error probability by choosing the input to be a random code from an i.i.d. Gaussian ensemble.
The proof is very similar to the one provided in [3] for the proof of theorem 2 of that paper. As has been done in [3], we can show that

$$P_e(\rho) \leq P_{out}(R) + P(error,\, no\, outage)\,. \qquad (43)$$

Now using (39), we get

$$P_e(\rho) \dot{\leq} \rho^{-d_{out}((1+\frac{L-1}{T})r)} + P(error,\, no\, outage)\,. \qquad (44)$$

In the same way as in [3], we can show that for a given channel, the error probability is upper bounded by

$$P(error\,|\mathbf{H}) \leq \rho^{T\,r} \det(\mathbf{I} + \frac{\rho}{2}\mathcal{A}_T(\mathbf{H})\mathcal{A}_T^H(\mathbf{H}))^{-1}\,. \qquad (45)$$

Using (11) and results of Lemma 6, we then show that

$$\begin{aligned} P(error\,|\mathbf{H}) &\leq \rho^{T\,r} e^{-\ln\det(\mathbf{I}_{N_r(T+L-1)} + \frac{\rho}{2}\mathcal{A}_T(\mathbf{H})\mathcal{A}_T^H(\mathbf{H}))} \\ &\leq \rho^{T\,r} e^{-T\frac{1}{2\pi j}\oint\frac{dz}{z}\ln\det(\mathbf{I}_{N_r} + \frac{\rho}{2}\mathbf{H}(z)\mathbf{H}^\dagger(z))} \\ &\leq e^{Tc^1}\rho^{T\,r}\det(\mathbf{I} + \frac{\rho}{2}L\,\bar{\mathbf{H}}\bar{\mathbf{H}}^H)^{-T} \end{aligned} \qquad (46)$$

where $c^1 = p\ln(qL\gamma_{qL}) + c$. This bound depends on $\bar{\mathbf{H}}$ only through the p non zero eigenvalues of $\bar{\mathbf{H}}\bar{\mathbf{H}}^H$, λ_i, $i = 1, \ldots, p$ sorted in a nondecreasing order. With the following change of variables $\lambda_i \triangleq \rho^{-\alpha_i}$, we then get

$$P(error\,|\alpha) \dot{\leq} \rho^{-T[\sum_{i=1}^p (1-\alpha_i)^+ - r]}\,. \qquad (47)$$

From the upper bound in (11) we can conclude that the *no outage* event is included in the event $(\mathcal{A})^c$, complement of the event $(\mathcal{A})$: $\{J(\mathbf{H}) < R\} = \{\alpha_1 \geq \ldots \geq \alpha_p \geq 0$, and $\sum_{i=1}^p (1-\alpha_i)^+ < r\}$,

$$P(error,\, no\, outage) \dot{\leq} \int_{(\mathcal{A})^c} p(\alpha)\,\rho^{-T[\sum_{i=1}^p (1-\alpha_i)^+ - r]}\,. \qquad (48)$$

The remainder of the proof is the same as the one in [3]. For $T \geq Lq + p - 1$ it leads to

$$P(error,\, no\, outage) \dot{\leq} \rho^{-d_{out}(r)}\,. \qquad (49)$$

Finally

$$\begin{aligned} P_e(\rho) &\dot{\leq} \rho^{-d_{out}((1+\frac{L-1}{T})r)} + \rho^{-d_{out}(r)} \\ &\dot{\leq} \rho^{-d_{out}((1+\frac{L-1}{T})r)}\,, \end{aligned} \qquad (50)$$

and $d^*(r) \geq d_{out}((1+\frac{L-1}{T})r)$. This completes the proof.

3. OPTIMALITY OF THE MMSE ZF DFE EQUALIZER FOR SIMO CHANNELS

For a SIMO channel ($N_t = 1$) the optimal tradeoff curve is the linear function $d^*(r) = LN_r(1-r)$, $0 \leq r \leq 1$. In [1] we show that the SNR at the output of the Decision Feedback Equalizer with a MMSE Zero-Forcing design verifies $\text{SNR}_{DFE} \geq \frac{\rho\,\|\mathbf{h}\|^2}{\gamma_L}$, $\|\mathbf{h}\|^2 = \sum_{i=0}^{L-1}\|\mathbf{h}_i\|_2^2$. We assume the use of a uniform QAM constellation of size ρ^r (corresponding to a rate $r\ln\rho$), and minimum distance $d^2 = \frac{3\sigma_x^2}{2(\rho^r-1)}$.
The transmitted frame is of size T and no space time coding is used. The transmitted symbols are detected sequentially at the receiver using the MMSE ZF DFE, which reproduces the same situation as the detection over a SISO flat channel with a SNR $=$ SNR_{DFE}, and this at each step. P_e can then be upper-bounded by $P_e \leq T\,P(E)$, where $P(E)$ is the detection error probability over the equivalent SISO flat channel. A point in the QAM constellation has at most four nearest neighbors, the $P(E)$ can then be upper-bounded by four times the Pairwise Error Probability with minimum distance. Finally using the PEP expression in[1] we get

$$P_e \leq 4T\,(1 + \frac{3}{8c_L(\rho^r-1)}\rho)^{-LN_r} \doteq \rho^{-LN_r(1-r)} = \rho^{-d^*(r)}\,. \qquad (51)$$

This shows the optimality of the MMSE ZF DFE equalizer for SIMO channels.

4. REFERENCES

[1] A. Medles and D.T.M. Slock. Decision-Feedback Equalization Achieves Full Diversity for Finite Delay Spread Channels. In *IEEE ISIT 2004*, Chicago, Illinois, Jun-Jul 2004.

[2] L. Grokop and D.N.C. TSe. Diversity/Multiplexing Tradeoff in ISI channels. In *IEEE ISIT 2004*, Chicago, Illinois, Jun-Jul 2004.

[3] L. Zheng and D. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Trans. Info. Theory*, 49(5):1073 –1096, May 2003.

[4] R. Gray. *Toeplitz and circulant matrices*. Stanford University, Palo Alto, 1977.

[5] 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.