

MIMO Capacity Pre-Log for Flat Fading Stationary Channels with no CSIR

Umer Salim

Mobile Communications Department
Eurecom
06560 Sophia Antipolis, France
umer.salim@eurecom.fr

Dirk Slock

Mobile Communications Department
Eurecom
06560 Sophia Antipolis, France
dirk.slock@eurecom.fr

Abstract—The asymptotic capacity for the non-coherent MIMO stationary channels having n_t transmit and n_r receive antennas with flat fading is the focus of this paper. Fading processes of concern are bandlimited. These non-coherent MIMO channels were studied by Etkin and Tse and lower bound of capacity was shown to grow with $\min(n_t, n_r)[1 - \min(n_t, n_r)\mu] \log(\text{SNR})$ where μ is the normalized Doppler bandwidth. The contribution of this paper is to specify the pre-log for MIMO channels and this is done by giving matching upper and lower bounds of the pre-log. Moreover $\min(n_t, n_r)$ factor in the pre-log term bears a small modification. The actual pre-log is $\min(n_t, n_r, \frac{1}{2\mu})[1 - \min(n_t, n_r, \frac{1}{2\mu})\mu]$ and takes into account the optimal number of streams that should be activated as a function of the Doppler bandwidth.

I. INTRODUCTION

Information theoretic bounds for various types of channels have got utmost importance since the explosion of research in MIMO promised new dimensions for data communication. Such capacity bounds are very important in the sense that they give the theoretical limits and motivate the researchers to approach these limits in practical systems. The area of capacity analysis for non-coherent (no CSIR and no CSIT) fading channels has received considerable attention in recent years since the usual assumption of perfect CSIR is not true in practical systems and channel realizations need to be estimated for correct decoding of data.

Usually block fading models are assumed for obtaining capacity bounds in the no CSIR case. In the standard version of this model [1], the fading remains constant over blocks consisting of T symbol periods, and changes independently from block to block. Capacity bounds are obtained by introducing training segments in an ad hoc fashion. For the standard block fading model, the capacity was shown [1], [2] to grow logarithmically with SNR. Later Liang and Veeravalli [3] allowed the fading to vary inside the block with a certain correlation matrix characterized by the rank Q and showed for SISO channels that the capacity pre-log is $(1 - Q/T)$.

Non-coherent capacity has also been analyzed with the channel fading process being symbol-by-symbol stationary. In this model, the fading is not independent but time selective without any block structure. Surprisingly, this model leads to very different capacity results: contrary to $\log(\text{SNR})$ capacity growth in block fading channels, here the capacity grows only

double logarithmically with SNR at high SNR [4], [5], [6] when the fading process is non-bandlimited, i.e. the channel prediction error is non-zero even if infinite past is known.

For symbol-by-symbol stationary Gaussian fading channels, if the Doppler spectrum is bandlimited, then the fading process is called non-regular and the prediction error given the infinite past goes to zero. Lapidoth [7] studied the SISO case for this kind of fading processes showing that the capacity grows logarithmically with SNR and the capacity pre-log is the Lebesgue measure of the frequencies where the spectral density of the fading process (Doppler spectrum) has nulls.

Etkin and Tse [8] study the same channel model of bandlimited fading for MIMO systems, they show that the pre-log exists even for MIMO systems with no CSIR but they only give a lower bound of the capacity pre-log.

In this paper, we specify the exact pre-log for such non-coherent MIMO channel giving matching upper and lower bounds. We should emphasize that the fading processes of interest to us in this paper are stationary and strictly bandlimited, the ones for which Lapidoth [7] established the capacity pre-log in the SISO case. The rest of the paper is organized as follows. In section II, we give the system model. Section III gives the main contribution of this paper. In section IV, we prove the upper bound of the MIMO capacity pre-log. In section V, we characterize the optimal number of active streams in terms of the Doppler bandwidth and complete the proof of the pre-log. Section VI gives the concluding remarks and also points to some future research directions.

II. SYSTEM MODEL

We consider a MIMO fading channel whose time- k output $Y[k] \in \mathbb{C}^{n_r}$ is given by

$$Y[k] = \sqrt{\frac{\text{SNR}}{n_t}} H[k] X[k] + Z[k] \quad (1)$$

where $X[k] \in \mathbb{C}^{n_t}$ denotes the time- k channel input vector, the matrix $H[k] \in \mathbb{C}^{n_r \times n_t}$ represents the time- k fading matrix consisting of i.i.d. circularly symmetric complex Gaussian components of zero mean and unit variance and $Z[k] \in \mathbb{C}^{n_r}$ denotes the additive Gaussian noise vector. Here \mathbb{C} denotes the complex field, n_t and n_r represent the number of transmit

and receive antennas respectively. We assume that the zero-mean circularly symmetric complex Gaussian noise is spatiotemporally white with spatial covariance matrix I_{n_r} , which represents the $n_r \times n_r$ identity matrix. The channel fading process $\{H[k]\}$ is assumed to be stationary, ergodic and with finite second order moment, i.e. $\mathbb{E}[|H[k]|^2] < \infty$. We take the fading process to be strictly bandlimited, so it is a non-regular stochastic process with limited Doppler spectrum support. The Lebesgue measure of the Doppler bandwidth is μ for each channel entry.

An average power constraint is imposed on the input hence

$$\mathbb{E}[|X[k]|^2] \leq n_t \quad (2)$$

Throughout this paper, $(\cdot)^T$ and $(\cdot)^\dagger$ will denote transpose and Hermitian transpose operators respectively.

The capacity pre-log is normally defined as

$$\text{PreLog} = \lim_{\text{SNR} \rightarrow \infty} \frac{C(\text{SNR})}{\log(\text{SNR})} \quad (3)$$

III. THE MAIN RESULT OF THE PAPER

The main result of this paper is the following theorem which characterizes the capacity pre-log for MIMO channels having Gaussian fading with bandlimited Doppler spectrum.

Theorem 1

$$\text{PreLog} = \min(n_t, n_r, \frac{1}{2\mu}) [1 - \min(n_t, n_r, \frac{1}{2\mu})\mu] \quad (4)$$

A. Outline of the Proof

For our MIMO system (1), the capacity C can be calculated from the following expression [9]

$$C = \lim_{n \rightarrow \infty} \frac{1}{n} \sup_{p(X^n)} I(X^{1:n}; Y^{1:n}) \quad (5)$$

where the maximization is done over all the input distributions which satisfy the imposed average power constraint (2). I represents the mutual information and we use $X^{1:n}$ as a shorthand for $(X[1], X[2], \dots, X[n])$.

To prove the given pre-log as the exact pre-log, we bound the mutual information (MI) from above and below and show that in both cases we achieve the same pre-log, hence proving theorem 1.

B. Capacity Upper Bound

The derivation of the upper bound is somewhat involved. First we represent the mutual information between the input and the output of our MIMO channel as the sum of two terms, the first of which is the mutual information when the channel information is known at the receiver and the second is the term which represents the loss due to not knowing the channel. Now the mutual information with CSIR known is a well studied case in the literature of information theory so results are readily available for this term and our task boils down to lower bounding the mutual information loss for not knowing the channel. We lower bound this loss and get our pre-log by combining with the pre-log of the mutual information with CSIR known.

C. Capacity Lower Bound

The derivation of the capacity lower bound is similar to as given by Etkin and Tse in [8] except that the number of the active transmit antennas depends also upon the Doppler bandwidth and the factor $\min(n_t, n_r)$ gets replaced by $\min(n_t, n_r, \frac{1}{2\mu})$. But in principle the derivation follows the same steps, hence we use their result because of space limitation and specialize it to the correct pre-log where minimization involves three factors n_t, n_r and μ . The main idea for this lower bound was to split the MI and then bound its corresponding components.

$$\text{MI} = I(X^{1:n}; Y^{1:n}) = h(Y^{1:n}) - h(Y^{1:n}|X^{1:n}) \quad (6)$$

The above equation follows from the definition of the MI [9] where $h(\cdot)$ represents the differential entropy. For lower bounding the capacity, any input can be selected and the resultant MI is a capacity lower bound hence in [8] they select i.i.d. Gaussian distribution as the input distribution. Using this input, they lower bound the entropy of the output and upper bound the conditional entropy of the output given the input achieving the lower bound of the pre-log as given by

$$\text{PreLog} \geq \min(n_t, n_r) [1 - \min(n_t, n_r)\mu] \quad (7)$$

IV. PROOF OF THE UPPER BOUND OF CAPACITY

Mutual information between the input and the output of (1) combined over n symbol times can be written as

$$\begin{aligned} \text{MI} &= I(X^{1:n}; Y^{1:n}) \\ &= I(X^{1:n}; Y^{1:n}, H^{1:n}) - I(X^{1:n}; H^{1:n}|Y^{1:n}) \end{aligned}$$

Above we used the chain rule of MI [9] to bring in the channel matrix. Both of the terms in the above equation can be split further by multiple application of the chain rule

$$I(X^{1:n}; Y^{1:n}, H^{1:n}) = I(X^{1:n}; H^{1:n}) + I(X^{1:n}; Y^{1:n}|H^{1:n}) \quad (8)$$

$$\begin{aligned} &I(X^{1:n}; H^{1:n}|Y^{1:n}) \\ &= I(X^{1:n}, Y^{1:n}; H^{1:n}) - I(H^{1:n}; Y^{1:n}) \\ &= I(X^{1:n}; H^{1:n}) + I(Y^{1:n}; H^{1:n}|X^{1:n}) - I(H^{1:n}; Y^{1:n}) \end{aligned}$$

combining the MI of the above two terms we get,

$$\begin{aligned} \text{MI} &= \\ &I(X^{1:n}; Y^{1:n}|H^{1:n}) - I(Y^{1:n}; H^{1:n}|X^{1:n}) + I(H^{1:n}; Y^{1:n}) \end{aligned} \quad (9)$$

where the first term is the MI with CSIR known and the next two terms will give some overall negative contribution representing the loss due to not knowing the channel. In the sub-sections below, we upper bound the MI (9) by bounding each term individually.

A. The First Term in Equation (9)

From Telatar's result [10], we know that the MI between the input and the output of a MIMO channel with CSIR known is given by the following expression

$$I(X^{1:n}; Y^{1:n}|H^{1:n}) = \min(n_t, n_r) n \log(\text{SNR}) + o(1) \quad (10)$$

where the $o(1)$ term becomes negligible when SNR goes to infinity.

B. The Second Term in Equation (9)

So now we need to lower bound the loss in MI due to not knowing the channel information at the receiver. This loss consists of the difference of two terms as shown in equation (9), and we bound each of these terms separately. Treating the first term of this loss, we get

$$\begin{aligned} & I(Y^{1:n}; H^{1:n} | X^{1:n}) \\ & \stackrel{a}{=} \mathbb{E}_{X^{1:n}} [h(Y^{1:n} | X^{1:n} = x^{1:n}) - h(Y^{1:n} | H^{1:n}, X^{1:n} = x^{1:n})] \\ & \stackrel{b}{=} \mathbb{E}_{X^{1:n}} [h(Y^{1:n} | X^{1:n} = x^{1:n}) - h(Z^{1:n})] \\ & \stackrel{c}{=} \mathbb{E}_{X^{1:n}} h(Y^{1:n} | X^{1:n} = x^{1:n}) - nn_r \log(\pi e) \end{aligned}$$

where (a) is by the definition of the MI as the difference of two entropy terms, (b) follows because the differential entropy is invariant to deterministic translations and in (c) we use the expression for the entropy of a multi-dimensional Gaussian distributed random variable [9].

Given the input, the output at each receive antenna is i.i.d. This fact was already shown in [8], moreover the output given the input ($Y|X$) is Gaussian distributed hence

$$\begin{aligned} & \mathbb{E}_{X^{1:n}} [h(Y^{1:n} | X^{1:n} = x^{1:n})] \\ & \stackrel{a}{=} n_r \mathbb{E}_{X^{1:n}} [h(Y_{n_r=1}^{1:n} | X^{1:n} = x^{1:n})] \\ & \stackrel{b}{=} n_r \mathbb{E}_{X^{1:n}} \log[\det \pi e R_{Y_{n_r=1}^{1:n} | Y_{n_r=1}^{1:n}}] \end{aligned} \quad (11)$$

In (a), we used the fact that the output given the input is i.i.d. at each receive antenna and $Y_{n_r=1}^{1:n}$ represents the signal received at first receive antenna ($n_r = 1$) for n symbol times and (b) follows by using the expression for the entropy of a multi-dimensional Gaussian distributed random variable [9] where $R_{Y_{n_r=1}^{1:n} | Y_{n_r=1}^{1:n}}$ is given by

$$R_{Y_{n_r=1}^{1:n} | Y_{n_r=1}^{1:n}} = \mathbb{E}[Y_{n_r=1}^{1:n} Y_{n_r=1}^{\dagger 1:n} | X^{1:n}] \quad (12)$$

For further processing, we change the system representation so as to focus on the signal received only at the first receiving antenna $n_r = 1$. Moreover this construction is such that we combine the input, the output and the corresponding channel coefficients (for first receive antenna only) over the duration of n symbol times. So the input signal is represented as a matrix of the following structure

$$\tilde{X} = \begin{bmatrix} X[1]^T & \mathbf{0}^T & \dots & \mathbf{0}^T \\ \mathbf{0}^T & X[2]^T & \ddots & \mathbf{0}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}^T & \mathbf{0}^T & \dots & X[n]^T \end{bmatrix} \quad (13)$$

hence the input matrix for n symbol times, \tilde{X} is a block diagonal matrix of size $n \times nn_t$. Moreover we take

$$\tilde{Y} = [y_1[1] \ y_1[2] \ \dots \ y_1[n]]^T \quad (14)$$

$$\tilde{Z} = [z_1[1] \ z_1[2] \ \dots \ z_1[n]]^T \quad (15)$$

$$\tilde{h} = [h_1[1]^T \ h_1[2]^T \ \dots \ h_1[n]^T]^T \quad (16)$$

where \tilde{Y} and \tilde{Z} are the column vectors of length n representing the vectors of the output and the noise samples at first receive

antenna for n symbol times and \tilde{h} is a vector of length nn_t where $h_1[k]^T$ represents the first row of the channel matrix $H[k]$, corresponding to n_t coefficients related to $n_r = 1$. Hence the equation for the system model becomes

$$\tilde{Y} = \sqrt{\frac{\text{SNR}}{n_t}} \tilde{X} \tilde{h} + \tilde{Z} \quad (17)$$

Using the new system representation, the covariance matrix of equation (12) can be written as

$$R_y = R_{Y_{n_r=1}^{1:n} | Y_{n_r=1}^{1:n}} = \mathbb{E}[\tilde{Y} \tilde{Y}^\dagger | \tilde{X}] \quad (18)$$

Using the value of \tilde{Y} from equation (17), the above covariance matrix becomes

$$\begin{aligned} R_y &= \frac{\text{SNR}}{n_t} \tilde{X} \mathbb{E}[\tilde{h} \tilde{h}^\dagger] \tilde{X}^\dagger + \mathbb{E}[\tilde{Z} \tilde{Z}^\dagger] \\ &= \frac{\text{SNR}}{n_t} \tilde{X} K_h \tilde{X}^\dagger + I_n \end{aligned} \quad (19)$$

In the above equations we used the fact that the noise samples are independent of the channel coefficients and the input, further they are i.i.d. spatiotemporally and K_h is the covariance matrix of \tilde{h} , hence of size $nn_t \times nn_t$.

For the product matrix $\tilde{X} K_h \tilde{X}^\dagger$, the rank will be mainly depending upon the rank of K_h as \tilde{X} with size $n \times nn_t$ will be of full rank n almost surely due to its block diagonal structure. \tilde{X} can be of reduced rank if and only if at a particular symbol time, symbols transmitted from all n_t antennas are deterministically zero. This can be avoided easily by putting some constraint on the code book of the source. This statement is also supported by Theorem 6 in [8]. So we can state that

$$\text{rank}(\tilde{X}) = n \quad \text{almost surely} \quad (20)$$

The covariance matrix K_h is surely a reduced rank matrix because the channel coefficients, although i.i.d. in space, have temporal correlation and the Lebesgue measure corresponding to the Doppler bandwidth is μ for each channel entry, hence the limiting rank of K_h when its size is made to go to infinity is given by

$$\lim_{n \rightarrow \infty} \text{rank}(K_h) = nn_t \mu \quad (21)$$

At this point we make another assumption that $n_t \mu \leq 1$, i.e. the product of the number of transmit antennas and the measure of the Doppler bandwidth per channel entry is less than or equal to 1, this point becomes clearer in section V which will indicate that if the Doppler spread is large, i.e. if the channel is varying quite rapidly, one might need to turn off some of the transmitting antennas to achieve optimal rates.

With this, the rank of $\tilde{X} K_h \tilde{X}^\dagger$ becomes

$$\text{rank}(\tilde{X} K_h \tilde{X}^\dagger) = nn_t \mu \quad (22)$$

Now the differential entropy of the output given the input of equation (11) can be written as

$$\begin{aligned} \mathbb{E}_{X^{1:n}} [h(Y^{1:n} | X^{1:n} = x^{1:n})] &= n_r \mathbb{E}_{X^{1:n}} \log[\det \pi e R_y] \\ &= n_r \mathbb{E}_{X^{1:n}} \log[\det \pi e (\frac{\text{SNR}}{n_t} \tilde{X} K_h \tilde{X}^\dagger + I_n)] \end{aligned} \quad (23)$$

In the above expression, the matrix $\tilde{X}K_h\tilde{X}^\dagger$ is a hermitian matrix, hence it can be diagonalized using eigen value decomposition (EVD) $\tilde{X}K_h\tilde{X}^\dagger = U\Lambda U^\dagger$, where U is a unitary matrix and Λ is a diagonal matrix of eigen values, where the number of non-zero eigen values is governed solely by the rank of K_h as previously explained. So the above equation can be written as

$$\begin{aligned} & \mathbb{E}_{X^{1:n}} h(Y^{1:n}|X^{1:n} = x^{1:n}) \\ &= n_r \mathbb{E}_{X^{1:n}} \log \det \left(\frac{\text{SNR}}{n_t} \Lambda + I_n \right) + nn_r \log(\pi e) \\ &= n_r \mathbb{E}_{X^{1:n}} \log \prod_{i=1}^{nn_t\mu} \left(\frac{\text{SNR}}{n_t} \lambda_i + 1 \right) + nn_r \log(\pi e) \\ &= n_r \mathbb{E}_{X^{1:n}} \sum_{i=1}^{nn_t\mu} \log \left(\frac{\text{SNR}}{n_t} \lambda_i + 1 \right) + nn_r \log(\pi e) \\ &= nn_r n_t \mu \log(\text{SNR}) + n_r \mathbb{E}_{X^{1:n}} \sum_{i=1}^{nn_t\mu} \log \left(\frac{\lambda_i}{n_t} + \frac{1}{\text{SNR}} \right) \\ &+ nn_r \log(\pi e) \end{aligned}$$

In the set of equations above, first we replace $\tilde{X}K_h\tilde{X}^\dagger$ by its EVD, then the determinant of the resultant diagonal matrix is written as the product of the diagonal elements, it becomes sum of the logarithm of these values and in the end we separate out the $\log(\text{SNR})$ factor from other terms.

Putting this value of the differential entropy in the MI term of equation (11), we get

$$\begin{aligned} & I(Y^{1:n}; H^{1:n}|X^{1:n}) \\ &= \mathbb{E}_{X^{1:n}} h(Y^{1:n}|X^{1:n} = x^{1:n}) - nn_r \log(\pi e) \\ &= nn_r n_t \mu \log(\text{SNR}) + n_r \mathbb{E}_{X^{1:n}} \sum_{i=1}^{nn_t\mu} \log \left(\frac{\lambda_i}{n_t} + \frac{1}{\text{SNR}} \right) \end{aligned} \quad (24)$$

C. The Third Term in Equation (9)

The term $I(H^{1:n}; Y^{1:n})$ represents the mutual information over a fictitious channel with input H , output Y and X plays the role of unknown channel fading (1). We should keep in mind that X is in fact the input and hence will be an unpredictable (regular) process with finite entropy rate. We do the analysis of $I(H^{1:n}; Y^{1:n})$ for two cases of $n_r < n_t$ and $n_r \geq n_t$ separately.

i) MI For $\mathbf{n}_r < \mathbf{n}_t$

For this case, the MI between the fictitious input (H) and the output Y has no growth with $\log(\text{SNR})$. Because fictitious channel process (X) introduces n_t parameters each symbol interval and with less number of receive antennas even this fictitious channel can not be resolved. Hence this MI falls under the category of non-coherent communication over regular fading channel process and gives no growth with $\log(\text{SNR})$ (see [4] for details).

It can also be shown by combining the MI of the first and the third terms

$$\begin{aligned} & I(X^{1:n}; Y^{1:n}|H^{1:n}) + I(H^{1:n}; Y^{1:n}) \\ &= [h(Y^{1:n}|H^{1:n}) - h(Y^{1:n}|H^{1:n}, X^{1:n})] \\ &+ [h(Y^{1:n}) - h(Y^{1:n}|H^{1:n})] \\ &= -h(Y^{1:n}|H^{1:n}, X^{1:n}) + h(Y^{1:n}) \end{aligned}$$

and this combined term has growth with $n_r \log(\text{SNR})$ per symbol time as shown below.

$$\begin{aligned} & h(Y^{1:n}) \leq nn_r \log(\text{SNR}) = n \min(n_t, n_r) \log(\text{SNR}) \\ & h(Y^{1:n}|H^{1:n}, X^{1:n}) = h(Z^{1:n}) = nn_r \log(\pi e) \end{aligned} \quad (25)$$

where the first inequality uses the independent bound and the Gaussian bound for the entropy of the output, combined with some high SNR approximation of neglecting the lower order terms. In the second equation we use the invariability of the differential entropy under deterministic translations and the expression for the entropy of the multi-dimensional Gaussian distributed random variable. This shows that per symbol time growth rate of the first term combined with the third term of the MI in equation (9) is $\min(n_t, n_r) \log(\text{SNR})$, but the same growth rate is obtained if we analyze the first term alone, as proved in section IV-A, or in other words the term $I(H^{1:n}; Y^{1:n})$ itself has no growth with $\log(\text{SNR})$ when n_r is less than n_t .

ii) MI For $\mathbf{n}_r \geq \mathbf{n}_t$

For the case when n_r exceeds n_t , although the fictitious channel (X) is still unpredictable over different time instants from the past observations but in each symbol interval the number of observations n_r is more than the number of fictitious channel parameters n_t . Hence n_t out of n_r observations get used to determine the fictitious channel parameters and still $(n_r - n_t)$ are left where the fictitious input H can be coherently detected. Now each of the receive antennas receives one row vector (n_t variables) of H (1) but each entry of H is of reduced bandwidth μ . So each of the excess receive antennas $(n_r - n_t)$ is able to resolve μn_t dimensions of the fictitious input H coherently which gives rise to growth with $\log(\text{SNR})$. This gives us the high SNR approximation as

$$I(H^{1:n}; Y^{1:n}) \leq n \mu n_t (n_r - n_t) \log(\text{SNR}) \quad \text{for } n_r \geq n_t \quad (26)$$

D. Result

We bounded each term of the mutual information expression (9) separately in the previous sub-sections. So the upper bound of this MI can be obtained by combining expressions from (10), (24), (25) and (26). Our point of concern in this contribution is the capacity pre-log, hence we divide the MI by $\log(\text{SNR})$ with limiting value of SNR going to infinity. Thus all the finite terms and lower order terms in the sum disappear. We divide this further by n to get the pre-log per symbol time and the resulting value is

$$\begin{aligned} \text{PreLog} &= \lim_{\text{SNR} \rightarrow \infty} \frac{I(X^{1:n}; Y^{1:n})}{n \log(\text{SNR})} \\ \text{PreLog} &\leq \begin{cases} \min(n_r, n_t) - \mu n_r n_t & n_r < n_t \\ \min(n_r, n_t) - \mu n_t n_t & n_r \geq n_t \end{cases} \end{aligned}$$

For the case when $n_t \leq n_r$ and $\min(n_t, n_r) = n_t$, the pre-log is $\min(n_t, n_r)[1 - \mu \min(n_t, n_r)]$. But when $n_r < n_t$ giving $\min(n_t, n_r) = n_r$, the pre-log is $\min(n_t, n_r)[1 - \mu n_t]$. Now there is this loss factor of $n_t \mu$ which clearly indicates that at very high SNR, the number of active transmitting antennas

should be selected no more than the receive antennas n_r as transmit antennas more than receive antennas do not add to the multiplexing gain $\min(n_t, n_r)$ of the system but they do increase the number of the channel coefficients to be estimated at each receive antenna and hence the loss factor. So by selecting $n_t = \min(n_t, n_r)$, the pre-log is upper bounded for both cases by

$$\text{PreLog} \leq \min(n_t, n_r)[1 - \min(n_t, n_r)\mu] \quad (27)$$

Combining the upper and the lower bounds from equations (7) and (27), the pre-log of the system is given by

$$\text{PreLog} = \min(n_t, n_r)[1 - \min(n_t, n_r)\mu] \quad (28)$$

V. LARGE DOPPLER BANDWIDTH ANALYSIS

Here we treat the case when the channel is still underspread but the inverse of the normalized Doppler bandwidth $1/\mu$ is comparable to $\min(n_t, n_r)$. Our expression of the pre-log equation (28) shows that the high SNR degrees of freedom depend entirely on $\min(n_t, n_r)$ and not on the individual values of n_t and n_r . As $n_t > n_r$ is strictly sub-optimal in high SNR non-coherent regime so let's take $n_t < n_r$ and the pre-log becomes $n_t(1 - n_t\mu)$. We want to analyze what is the optimal value of n_t for a fixed n_r and Doppler bandwidth μ . Figure (1) shows that for large enough n_r , initially pre-

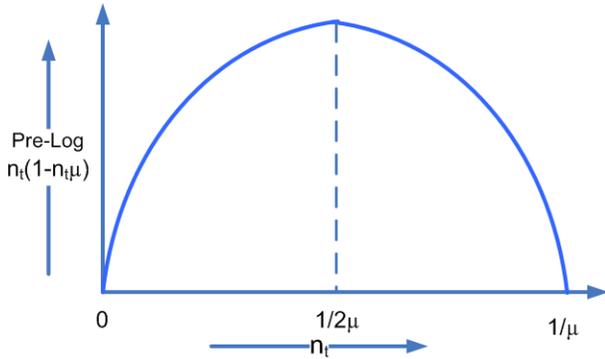


Fig. 1. Pre-Log versus Active Number of Streams

log increases with increasing n_t reaching its maximum value at $1/(2\mu)$ and starts decreasing onwards. The explanation is that n_t factor represents the number of independent streams which one can multiplex over this system but the coherent reception of this number of streams first requires estimation of the corresponding channel coefficients hence the loss factor also increases with the factor n_t . Now with large Doppler bandwidth when the coherence time is very short, using more streams means a greater loss factor which is proportional to the Doppler bandwidth. But due to very short coherence time, the coherent transmission does not last long enough to compensate that loss factor and to reduce the number of active streams becomes the optimal strategy. Hence if $n_t > 1/(2\mu)$, the active number of transmit antennas should be reduced to $1/(2\mu)$. This discussion indicates that the active number of transmit antennas (streams) in non-coherent MIMO should actually be

$\min(n_t, n_r, \frac{1}{2\mu})$ and the pre-log for a non-coherent MIMO system becomes

$$\text{PreLog} = \min(n_t, n_r, \frac{1}{2\mu})[1 - \min(n_t, n_r, \frac{1}{2\mu})\mu] \quad (29)$$

VI. CONCLUDING REMARKS

We have shown the exact capacity pre-log for Gaussian MIMO channels with bandlimited fading. We also characterized the optimal number of active transmit antennas (streams) in terms of the Doppler bandwidth of the channel fading process to achieve the optimal pre-log at high values of SNR.

The analysis of the upper bound of the mutual information shows that to characterize the capacity more precisely, the knowledge of the capacity achieving distribution may play vital role. So investigation for the optimal or close to optimal input distributions could be a fertile area of research as it will make possible the evaluation of the lower order terms and the constants which appear besides $\log(\text{SNR})$.

ACKNOWLEDGMENT

The authors would like to acknowledge helpful discussions and comments by Randa Zakhour.

Eurecom's research is partially supported by its industrial members: BMW Group Research & Technology, Bouygues Telecom, Cisco, Hitachi, ORANGE, SFR, Sharp, STMicroelectronics, Swisscom, Thales. The research work leading to this paper has also been partially supported by the European Commission under the ICT research network of excellence NewCom++ of the 7th Framework programme and by the French ANR project APOGEE.

REFERENCES

- [1] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communications link in rayleigh flat fading," *IEEE Trans. on Information Theory*, vol. 45, pp. 139–157, January 1999.
- [2] L. Zheng and D. N. C. Tse, "Communication on the grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. on Information Theory*, vol. 48, pp. 359–383, February 2002.
- [3] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective rayleigh-fading channels," *IEEE Trans. on Information Theory*, vol. 50, pp. 3095–3110, December 2004.
- [4] A. Lapidoth and S. Moser, "Capacity bounds via duality with applications to multi-antenna systems on flat fading channels," *IEEE Trans. on Information Theory*, vol. 49, pp. 2426–2467, October 2003.
- [5] T. Koch and A. Lapidoth, "The fading number and degrees of freedom in non-coherent mimo fading channels: a peace pipe," in *Proc. IEEE International Symposium on Information Theory*, September 2005, vol. 52, pp. 437–453.
- [6] A. Lapidoth and S. Moser, "The fading number of single-input multiple-output fading channels with memory," *IEEE Trans. on Information Theory*, vol. 52, pp. 437–453, February 2006.
- [7] A. Lapidoth, "On the asymptotic capacity of stationary gaussian fading channels," *IEEE Trans. on Information Theory*, vol. 51, pp. 437–446, February 2005.
- [8] R. Etkin and D. N. C. Tse, "Degrees of freedom in some underspread mimo fading channels," *IEEE Trans. on Information Theory*, vol. 52, pp. 1576–1608, April 2006.
- [9] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, New York: John Wiley and Sons, 1991.
- [10] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, pp. 585–595, November 1999.