Non-coherent Rayleigh fading MIMO channels: Capacity Supremum

Rasika R. Perera, Tony S. Pollock, Thushara D. Abhayapala
Department of Information Engineering,
Research School of Information Sciences and Engineering,
The Australian National University, ACT 0200, AUSTRALIA.
e-mail: rasika.perera@anu.edu.au

Abstract—This paper investigates the limits of information transfer over a fast Rayleigh fading MIMO channel, where neither the transmitter nor the receiver has the knowledge of the channel state information (CSI) except the fading statistics. We develop a scalar channel model due to absence of the phase information in non-coherent Rayleigh fading and derive a capacity supremum with the number of receive antennas at any signal to noise ratio (SNR) using Lagrange optimisation. Also, we conceptualise the discrete nature of the optimal input distribution by posing the optimisation on the channel mutual information for \( N \) discrete inputs. Furthermore, we derive an expression for the asymptotic capacity when the input power is large, and compare with the existing capacity results when the receiver is equipped with a large number of antennas.

Index Terms—Channel capacity, mutual information, Rayleigh fading, upper bound, SISO, MIMO, Lagrange optimisation.

I. INTRODUCTION

Communication over rapidly time-varying channels where the receiver is unable to estimate the channel state is a challenging task. In particular, for a mobile receiver the estimation may become difficult and the limits of information transfer in this scenario is vital when the channel becomes non-coherent. In this paper we study the capacity of the non-coherent Rayleigh fading MIMO channel to identify the limits of information transfer with no CSI at both ends.

One of the important problems in information theory is that of computing the capacity of a given communication channel, and finding the optimal input distribution that achieves capacity. The capacity problem is addressed by maximising the mutual information, a concave function subject to some constraints on the channel input distribution. For discrete finite-size input alphabets, this is a finite dimensional problem and all the input constraints are such that they define a compact, convex set, where the optimisation is straightforward. However, for continuous input channels, neither the continuity of the objective function, nor the compactness of the constraint set is now obvious, and the optimisation for capacity is considered a difficult problem [5], [8].

A large amount of literature has appeared in the area of multi-antenna wireless communications over the past decade. Much of the interest was motivated by the work in [1], [12] which shows a potential increase in the channel capacity using multiple antennas for both transmitting and receiving. The capacity of the MIMO channel, when neither the transmitter nor the receiver has CSI, has traditionally been considered an open and difficult problem. In this paper, we provide an upper bound on the non-coherent Rayleigh fading MIMO channel capacity for any receive antenna number at any SNR and elaborate on the discrete nature of the optimal input.

The MIMO channel capacity derived in [1], is based on the assumptions that i) the channel matrix elements are independent and circularly symmetric\(^\text{1}\) complex Gaussian; ii) the noise at different receive antennas are independent and white and; iii) the CSI is perfectly known at the receiver but not to the transmitter. The main result is linear capacity growth with the minimum number of transmit and receive antennas. Also [1] shows that the optimum input distribution which achieves this channel capacity is circularly symmetric complex Gaussian which maximises the channel output entropy when the CSI is unknown.

The capacity evaluation using Monte Carlo method when the receiver CSI is perfectly known is shown in [2]. The results extends Telatar’s work [1] and provides accurate numerical results. Furthermore [3] shows the use of water filling on the input Gaussian vector when receiver and transmitter CSI is available at the both transmitter and receiver. The increase in capacity is significant for small number of antennas and at low SNR. At high SNR the fading is less destructive and the water filling solution does not provide a considerable gain over the capacity with CSI at the receiver only.

The non-coherent channel capacity of the time selective Rayleigh fading channel is studied in [4]. Upper and lower bounds on the capacity at high SNR for single antenna systems is shown, and derived a lower bound on the capacity with multi-antennas. Abou-Faycal, Trott and Shamai [5] investigated the single input single output (SISO) discrete time memoryless Rayleigh fading channel, proving that the capacity achieving distribution is discrete with a finite number of mass points. In addition, Taricco [6] showed the capacity supremum and confirmed that the attainable input distribution is discrete in agreement with Abou-Faycal’s results. A general

\[ e^{j\theta} \]  

\( e^{j\theta} \) is identical to the distribution of \( \omega \).

\( \text{1}\) The distribution of a complex variable \( \omega \) is said to be circularly symmetric if for any deterministic \( -\pi \leq \theta \leq \pi \), the distribution of random variable 

---

1 T. S. Pollock and T. D. Abhayapala have appointments with National ICT Australia (NICTA). Parts of this work has been approved at Asia Pacific Conference on Communications (APCC), pp. 72-76, Oct. 2005, Perth, Australia.
channel model is considered in [7] ignoring the structure of the channel.

Conversely in [8], the channel is assumed to have only additive noise, and in [5], [6], the channel is assumed to have both additive and multiplicative noise. In [7] it is shown that under a peak power constrained input, the capacity achieving input distribution is discrete, and attempts to make a case for the argument that continuous capacity achieving distributions are actually anomalies, and for most channels the optimal input distribution is discrete with a finite number of mass points.

Similar work is reported by Marzetta and Hochward [9] in non-coherent Rayleigh fading MIMO channels using a block fading model over a coherence interval of $T$ symbol periods. They characterised a certain structure of the optimal input distribution and computed the channel capacity with reduced complexity. The coherence capacity with channel coherence time $T$ is used as the upper bound.

Also, it is shown that the non-coherent channel capacity approaches the coherent capacity as $T$ becomes large where the optimal input is approximately independent complex Gaussian. Zheng and Tse [10] extended this work and specifically computed the asymptotic capacity at both high and low SNR in terms of $T$ and the number of transmit and receive antenna elements. The suggested input is to transmit orthogonal vectors on a certain number of transmit antennas with constant equal norms. They also claim that having more transmit antennas than receive antennas does not provide any capacity gain at high SNR, while having more receive antennas yield a capacity gain.

Recently, Lapidoth and Moser [11] showed the capacity of non-coherent multi-antenna systems grows only double logarithmically in SNR and evaluated the fading number$^2$ with the optimum input distributions in the high SNR region. This double logarithmic behavior is a good example to visualise the low capacity available with no CSI compared to the coherent capacity given for MIMO [1], [12] and SISO [13] systems respectively.

In this paper we address the non-coherent uncorrelated Rayleigh fading MIMO channel and show a capacity supremum optimising the mutual information under average power constrained input for any SNR.

The contributions of the paper as follows:

1) We provide the mutual information of the non-coherent Rayleigh fading MIMO channel in simple form using output differential entropies.
2) We optimise the mutual information using Lagrange optimisation method and show a capacity supremum for a given number of receive antennas at any SNR.
3) We show the asymptotic analysis of the capacity supremum with double logarithmic behavior at high SNR, similar to the results shown in [11] and conjecture the discrete nature of the optimal input.
4) The proposed capacity in this paper can be taken as an upper bound$^3$ to the non-coherent uncorrelated MIMO channel capacity in Rayleigh fading.

The organisation of this paper is as follows. Section II contains the channel model with notations used for the non-coherent Rayleigh fading MIMO communication system. The derivation of mutual information for the introduced channel model is presented in Section III along with the detailed work based on Lagrange optimisation for channel capacity in IV. Section V presents the numerical results and analysis of our results. Finally conclusions are drawn in Section VI.

II. MIMO CHANNEL MODEL

The input output relationship of a MIMO channel can be written as

$$Y = HX + N,$$

(1)

where the output $Y$ is $n_r \times 1$, the channel gain matrix $H$ is $n_r \times n_t$. The input $X$ is $n_u \times 1$ and the noise $N$ which is assumed to be zero mean complex Gaussian is $n_u \times 1$. Each element of $H$, $h_{ij}$, $i = 1, \ldots, n_r$, $j = 1, \ldots, n_t$ is assumed to be zero mean circular complex Gaussian random variables with a unit variance in each dimension.

We use $X = \{X\}$ and $Y = \{Y\}$ to denote the random scalar variables where $|·|$ is the Euclidean norm. $x$ and $y$ represent each realisation of $X$ and $Y$ (i.e. $x \in X$ and $y \in Y$). The input is power limited with an average power constraint $\int x^2 p_X(x)dx \leq P$. $n_t$ and $n_r$ denote the number of transmit and receive antennas respectively, and $\gamma = -\int_0^\infty e^{-\nu}\log \nu d\nu \approx 0.5772, \ldots$ denotes the Euler’s constant. We use $\Gamma(·)$ and $\Psi(·)$ to indicate Gamma and Psi functions respectively. $h(X)$ denotes the differential entropy of $X$, and $I(X;Y)$ designates the mutual information between $X$ and $Y$. The expected value over a set of random variables are denoted by $E\{·\}$, with $\det(·)$ for the determinant, $(·)^*$ for conjugate transpose of a matrix and $I$ for an identity matrix. All the differential entropies and the mutual information are defined to the base “e”, and the results are expressed in “nats”. Neither the receiver nor the transmitter has the knowledge of CSI except the fading statistics. Channel coherence time is one where the channel changes independently at every transmitted symbol.

III. MIMO MUTUAL INFORMATION

A. Capacity with Receiver CSI

The Rayleigh fading MIMO channel capacity when the receiver has the perfect CSI, given by,

$$C_{\text{resi}} = E_H \left[ \log \det \left( I_{n_r} + \frac{P}{n_t} HH^* \right) \right]$$

(2)

was derived by Telatar in optimising the mutual information $I(X;Y) = h(Y) - h(N)$ [1] and later Foschini [12] who extended the work to show how the capacity scales with increasing SNR for a large but practical number of antenna elements at both the transmitter and receiver. The linear

$^2$The fading number is defined as the limit of the difference between the channel capacity and $\log(1 + \log(1 + \rho))$, where $\rho$ is the SNR.

$^3$The term “upper bound” is used since for any input distribution either discrete or continuous, the mutual information achieved through the channel is lower than the capacity result derived in this paper.
growth of the capacity with $n_r$ is shown for a special case where the channel matrix $H = I_{n_r,n_r}$. The capacity increase in [2] is more prominent having multiple antennas at the receiver instead of the transmitter. However, under fast fading conditions, the estimation of fading coefficients which are assumed to be independent could be difficult due to the short duration of fades. It is of interest to study the capacity of such a channel and understand the ultimate limits when no CSI is available. Furthermore, the increasing demand for higher data rates along with mobility will make the instantaneous channel measurement more difficult. Therefore, it is important to find the optimal rate when the CSI is not perfectly available at the receiver (non-coherent). In this paper, we consider the mutual information of non-coherent uncorrelated Rayleigh fading MIMO channel and investigate the capacity.

### B. Mutual Information

The conditional probability density function (pdf) of the output given input of channel model (1) is given by

$$f_{Y|X}(y|x) = \frac{1}{(2\pi(1+x^2))^{n_r}} \exp \left[-\frac{y^2}{2(1+x^2)}\right], \quad (3)$$

where $f_{Y|X}(y|x) \triangleq f_{Y|X}(y_1,y_2, \ldots, y_{n_r}|x_1,x_2, \ldots, x_{n_r})$, $y \in [Y]$ and $x \in [X]$. The magnitude sign is removed in (3) for simplicity and likewise in the rest of this paper since the non-coherent Rayleigh fading channel does not carry any phase information [6]. The pdf of the magnitude distribution of (1) has the form

$$p_Y|X(y|x) = \frac{y^{2n_r-1} \exp \left[-\frac{y^2}{2(1+x^2)}\right]}{2^{n_r-1} \Gamma(n_r)(1+x^2)^{n_r}}, \quad (4)$$

when Jacobian coordinate transformation is applied on $2n_r$ dimensions. The output conditional entropy $h(Y|X)$ for (1) is given by

$$h(Y|X) = -E_x \left\{ \int_0^\infty p_Y|X(y|x) \log [p_Y|X(y|x)] dy \right\}, \quad (5)$$

where the expectation is taken over $x \in X$. With (4), we get

$$h(Y|X) = \frac{1}{2} E_x \left\{ \log(1+x^2) \right\} + \log \left[ \frac{\Gamma(n_r)}{\sqrt{2}} \right] - \left(n_r - \frac{1}{2}\right) \Psi(n_r) + n_r. \quad (6)$$

Equation (6) can be used to calculate the output conditional entropy of uncorrelated Rayleigh fading MIMO channel when no CSI is available for a given input distribution. With the output entropy $h(Y) = -\int_0^\infty p_Y(y) \log[p_Y(y)]dy$, we obtain the mutual information [14]

$$I(X;Y) = h(Y) - h(Y|X) = -\int_0^\infty p_Y(y) \log[p_Y(y)]dy - \frac{1}{2} E_x \left\{ \log(1+x^2) \right\} - \log \left[ \frac{\Gamma(n_r)}{\sqrt{2}} \right] + \left(n_r - \frac{1}{2}\right) \Psi(n_r) - n_r. \quad (7)$$

From (7), the mutual information for a given input distribution can be computed. However, finding the optimal input and hence the capacity for a given input constraint is difficult. In next Section, we show how to derive an upper bound on (7) identifying some key properties of the optimal input.

### IV. Non-Coherent MIMO Capacity

#### A. Output Constraints

To obtain an expression for capacity of no-coherent Rayleigh fading MIMO channel, equation (7) needs to be maximised subject to an appropriate constraint on the input. Usually constraints used are $\int_0^\infty p_X(x)dx = 1$ and $\int_0^\infty x^2 p_X(x)dx = P$. However, the maximisation of (7) subject to these constraints does not provide a valid input pdf for SISO or MIMO in this case. To overcome this difficulty, additional constraints are used in [6] for SISO non-coherent Rayleigh fading channel. Likewise, to optimise (7) we use the following constraints

$$\int_0^\infty p_Y(y)dy = 1, \quad \int_0^\infty y^2 p_Y(y)dy = 2n_r(1+P), \quad \int_0^\infty p_Y(y)\log ydy = \frac{1}{2}(\beta + \Psi(n_r) + \log 2), \quad (8a,b,c)$$

where $\beta = E_x \left\{ \log(1+x^2) \right\}$. The second constraint is the average mean squared power of $y \epsilon Y$, which is considered as the induced power at the output by the input, channel gain and noise. The constraint (8c) is derived in Appendix VII-A. This additional constraint is used to support the optimisation process in order to arrive at a valid output pdf. Similar techniques are commonly employed in convex optimisation work [15].

#### B. Lagrange Optimisation

Using the Lagrange variable $L$ and the multipliers $\lambda_1$, $\lambda_2$, and $\lambda_3$, we define

$$L = I(X;Y) + \lambda_1 \left\{ \int_0^\infty p_Y(y)dy - 1 \right\} + \lambda_2 \left\{ \int_0^\infty y^2 p_Y(y)dy - 2n_r(1+P) \right\} + \lambda_3 \left[ \int_0^\infty p_Y(y)\log ydy - \frac{1}{2}(\beta + \Psi(n_r) + \log 2) \right]. \quad (9)$$

Solving (9) for $p_Y(y)$, we obtain the optimum output pdf

$$p_Y(y) = \exp \left[ \lambda_1 - 1 + \lambda_2 y^2 + \lambda_3 \log y \right] \quad (10)$$

for the mutual information in (7) in terms of Lagrange variables.

We substitute optimum $p_Y(y)$ (10) into three constraints (8a), (8b), and (8c) and use the integral identities [16, Page 360-365] to derive the following three equations:

$$\frac{(-\lambda_2)^{-1}(1+\lambda_3)}{2} e^{\lambda_1-1} \Gamma \left( \frac{1+\lambda_3}{2} \right) = 1, \quad (11a)$$
Equation (13) can be solved for
\[ \frac{(-\lambda_2)^{\frac{3+\lambda_3}{2}} e^{\lambda_1 - 1} \Gamma \left( \frac{3 + \lambda_3}{2} \right)}{2} = 2n_r(1 + P), \tag{11b} \]
\[ \frac{e^{\lambda_1 - 1}(-\lambda_2)^{\frac{1+\lambda_3}{2}}}{4\sqrt{-\lambda_2}} \left\{ \Psi \left( \frac{1 + \lambda_3}{2} \right) - \log(-\lambda_2) \right\} = \frac{1}{2}(\beta + \Psi(n_r) + \log 2), \tag{11c} \]
where the second Lagrange multiplier \( \lambda_2 \) is a negative quantity.

From (11a) and (11b), and using the relationship [17, Page 255]
\[ \Gamma \left( \frac{3 + \lambda_3}{2} \right) = \left( \frac{1 + \lambda_3}{2} \right)^\Gamma \left( \frac{1 + \lambda_3}{2} \right), \]
we get
\[ \lambda_3 = -4\lambda_2 n_r(1 + P) - 1. \tag{12} \]

With (12) and expressing the quantity \( e^{\lambda_1 - 1} \) in terms of \( \lambda_2 \), we can solve (11c) for \( \lambda_2 \) with
\[ \Psi[-2\lambda_2 n_r(1 + P)] - \log(-\lambda_2) = \beta + \Psi(n_r) + \log 2. \tag{13} \]
Equation (13) can be solved for \( \lambda_2 \) for certain \( n_r, P \) and \( \beta \) which is a function of \( P \). We assume, there exists a solution in the form
\[ \lambda_2 = \frac{-\zeta}{2n_r(1 + P)}, \quad \zeta > 0. \tag{14} \]
This assumption is made only to ease the rest of the mathematics involved and has no effect on the capacity results. From \( \lambda_2 \) we obtain
\[ \beta = \log \left\{ \frac{n_r(1 + P)}{\zeta} \right\} + \Psi(\zeta) - \Psi(n_r), \tag{15} \]
and following two equations for \( \lambda_3 \) and \( \lambda_1 \):
\[ \lambda_3 = 2\zeta - 1, \tag{16a} \]
\[ e^{\lambda_1 - 1} = \frac{\zeta^2}{n_r(1 + P) \Gamma(\zeta)}. \tag{16b} \]
Substituting these Lagrange multipliers in (16) we get the optimum output pdf
\[ p_Y(y) = \frac{\zeta^2 y^{2\zeta - 1}}{n_r(1 + P) \Gamma(\zeta)} \exp \left[ -\frac{\zeta y^2}{2n_r(1 + P)} \right]. \tag{17} \]

We have the following remarks:

1) To calculate capacity we need to find values for \( \lambda_1, \lambda_2 \) and \( \lambda_3 \).
2) Equation (13) could be used to calculate \( \lambda_2 \), for a given value of \( P \) and \( \beta \). However, it is difficult to find the value of \( \beta = E_x \{ \log(1 + x^2) \} \) for a given \( P \).
3) Note that \( 0 \leq \beta \leq \log(1 + P) \), where the right inequality is derived by direct application of Jensen’s inequality [18].
4) To better understand the characteristics of \( \beta \) we use (13) to plot both \( \beta \) and \( \beta - \log(1 + P) \) vs \( \zeta \) in Fig. 1. It is clear that \( \beta - \log(1 + P) \) is an increasing function of \( \zeta \).
5) In next Section we derive a supremum for the capacity using the properties of \( \beta \).

C. Capacity Supremum

Substituting the optimum \( p_Y(y) \) from (17) and \( \beta \) from (15) into (7), we obtain the non-coherent channel capacity
\[ C(\zeta) = G(\zeta) - G(n_r), \tag{18} \]
where
\[ G(\tau) = \log \Gamma(\tau) + \tau (1 - \Psi(\tau)). \tag{19} \]

From Fig. 1 it is clear that low \( n_r \) gives high \( \beta \) for a given SNR. However, \( \beta \geq 0 \) since the expectation over the function \( \log(1 + x^2) \) is always a positive quantity. Therefore, the \( \zeta \) which produce negative \( \beta \) values does not provide a valid capacity for a given \( n_r \). The capacity \( C(\zeta) \) is a monotonically decreasing function of \( \zeta \) since
\[ \frac{\partial C(\zeta)}{\partial \zeta} = -\zeta \Psi_1(\zeta), \tag{20} \]
where \( \Psi_n(\cdot) \) is the \( n \)-th derivative of \( \Psi(\cdot) \) [17, page 253-255]. Fig. 2 depicts the channel capacity for \( \zeta \in (0,1) \). The capacity can be computed for some \( \beta = \beta' \), seeking the optimal \( \zeta \) which satisfies the input power constraint. Furthermore, \( \beta \) in (13) is a monotonically increasing function of \( \zeta \) where
\[ \frac{\partial \beta(\zeta)}{\partial \zeta} = \frac{1}{\zeta} + \Psi_1(\zeta). \tag{21} \]
Therefore, the supremum of (18)
\[ C_{\text{sup}} = C(\zeta_s) = G(\zeta_s) - G(n_r) \tag{22} \]
is obtained with \( \beta = 0 \) where the corresponding \( \zeta_s \) is given by
\[ \Psi(\zeta_s) - \log(\zeta_s) = \Psi(n_r) - \log[n_r(1 + P)]. \tag{23} \]

The input power \( P \) in (23) vs \( \zeta_s \) as a function of \( n_r \) is given in Fig. 3. It is clear that there exist \( \zeta_s \) which gives the solution to (23) for any \( P \) for a given \( n_r \). Fig. 4 shows the capacity supremum in (22) against the input power for different \( n_r \) equating \( \zeta_s \) in (24). This non-coherent Rayleigh
fading MIMO channel capacity supremum can be used as an upper bound for all input distributions.

For larger \( n_r \), Sengupta and Mitra [19] show the capacity of non-coherent Rayleigh fading MIMO channel

\[
C_{\text{MIMO}} = \frac{1}{2} \log \left( \frac{n_r}{2\pi} \right) + \log \mu + \frac{P}{\mu n_t \left( 1 + \frac{P}{n_t} \right)},
\]

(24)

where

\[
\mu = \int_0^\infty \frac{dy}{1 + y} \exp \left[ -\frac{y}{\mu \left( 1 + \frac{P}{n_t} \right)} \right]
\]

and \( \mu \approx \log(1 + \frac{P}{n_t}) \) for large \( P \). Furthermore, they established a continuous input distribution which achieves (24) for large \( n_r \). We will discuss the analysis of the capacity supremum in (22) with respect to (24) in Section V.

The capacity (22) is independent of the number of transmit antennas since the optimisation is carried out using magnitude of the input vector. Therefore, the effect of number of transmit antennas on capacity is not apparent. However, [9] proves that the capacity of \( n_t > T \) is equal to the capacity for \( n_t = T \) where \( T \) is the channel coherence time. In this paper, we consider \( T = 1 \), and therefore according to the theorem 1 in [9], the optimal \( n_t = 1 \). Therefore, we conclude that capacity supremum (22) is true irrespective of \( n_t \).

### D. Optimal Input Distribution

The corresponding input distribution which provides this channel capacity supremum (22) for a certain \( n_r \) is given by

\[
\int_0^\infty p_X(x) \frac{y^{2n_r-1} \exp \left[ -\frac{y^2}{2(1+x^2)} \right]}{2^{n_r-1} \Gamma(n_r)(1+x^2)^{n_r}} \, dx = \frac{\zeta_s \zeta \mu^{2\zeta_s-1}}{n_r(1+P)^{\zeta_s} \Gamma(\zeta_s)} \times \exp \left[ -\frac{\zeta s y^2}{2n_r(1+P)} \right].
\]

(25)

The integral in (25) takes the form

\[
\int_a^b K(s,t)f(t) \, dt = g(s),
\]

(26)

a well known Fredholm equation of the first kind [20] where \( K(s,t) \) is the kernel. In general, such problems are ill-posed where small changes to the problem can make very large changes to the answer obtained. The kernel in (25) is analytic in \( y \) over the whole plane for any \( n_r \). However, the right hand side of (25) and its derivative with respect to \( y \) is infinite when \( y \rightarrow 0 \) for any \( n_r \) and \( \zeta_s \). Therefore, (25) does not provide a continuous solution for \( p_X(x) \) in which the \( C_{\text{sup}} \) in (22) is attained. This leads us to find a discrete input distribution in...
the form of
\[ p_X(x) = \sum_{i=1}^{N} p_i \delta(x - x_i), \]  
(27)

where \( p_i \) and \( x_i \) to be obtained solving
\[ g(s) = \sum_{i=1}^{N} p_i K(s, t_i). \]  
(28)

The number of transmitters \( n_t \) has no effect on the result. The only requirement is to Euclidean norm of the input vector \( X \) be discrete irrespective of transmit diversity. For a discrete input, we can pose a new optimisation problem to compute the channel capacity
\[
C_{\text{dis}} = \sup_{\sum_{i=1}^{N} p_i x_i^2 \leq P} \int_0^\infty \int_0^\infty p_X(x)p_Y|X(y|x) \times \log \frac{p_Y|X(y|x)}{p_X(x)p_Y|X(y|x)} \, dy \, dx \\
= \sup_{\sum_{i=1}^{N} p_i x_i^2 \leq P} \int_0^\infty \int_0^\infty p_X(x)p_Y|X(y|x) \times \log \frac{p_Y|X(y|x)}{\sum_{j=1}^{N} p_j p_Y|X(y|x_j)} \, dy, \]  
(29)

subject to the input power constraint \( P \). If the solution exists for (29), it will provide a good lower bound to \( C_{\sup} \). However, this optimisation problem is extremely difficult since the number of discrete points \( N \) is unknown and the optimum probabilities and their locations to be found satisfying the input power constraint. Similar work is reported in [5] for a single antenna case and the numerical evaluation is given using the Kuhn T Tucker conditions in order to verify the optimality of the capacity achieving mass point probabilities and their locations. Further work is required in this area to identify the optimal input at any SNR. The capacity supremum in (22) can be treated as an upper bound for the capacity of non-coherent Rayleigh fading MIMO channel.

E. Capacity for \( \beta > 0 \)

In addition to the capacity supremum found in the previous subsection, we now elaborate another solution when \( \beta > 0 \) as the capacity for a MIMO system with a small number of receive antennas \( n_r \). Since \( \beta \) is an increasing function of \( \zeta \) for a given \( n_r \), we can obtain the capacity
\[
C_{\sup}(\alpha) = G(\alpha) - G(n_r), \]  
(30)

for the input power
\[
P = \frac{\alpha}{2n_r} \exp \left[ 2(\Psi(n_r) + \log 2) - \Psi(\alpha) \right] - 1, \]  
(31)

when \( \beta = \Psi(n_r) + \log 2 > 0. \)

The quantity \( \alpha \in (0, \infty) \), and the upper limit of \( P \) is a function of \( n_r \). For large \( n_r \) and low input power there is no \( \alpha \) which provides a solution to (31). This could be seen from the asymptotic value of (31)
\[
P_{\text{asy}} = \frac{1}{2n_r} \exp \left[ 2(\Psi(n_r) + \log 2) \right] - 1, \]  
(32)

when \( \alpha \) approaches infinity since the minimum \( P \) is very high for \( n_r \geq 2 \). Fig. 5 illustrates this, where there is no solution for \( P = 0 \) in (31) when \( n_r \geq 2 \).

This excludes us in finding a solution to (30) for \( n_r \geq 2 \). However, we can show the capacity in (30) for \( n_r = 1 \) when the solution exists for \( \alpha \) in the whole input range of \( P \).

For \( n_r = 1 \), (22) and (28) becomes
\[
C_{\sup}(\zeta)_{n_r=1} = (\alpha - \gamma - 1) + \log \Gamma(\zeta) - \zeta \Psi(\zeta) \]  
(33)

and
\[
P_{\text{asy}} = \zeta e^{-\Psi(\zeta)} - 1 \]  
(34)
Multiplying the both sides of (39) by \( \log n \), we get
\[
\zeta_s \approx \log n + P
\]
and degrees of freedom defined by \( n_{\min} = \min\{n_t, n_r\} \). The asymptotic capacity
\[
C(\text{SNR}) = \log[1 + \log(1 + \text{SNR})] + \chi + o(1),
\]
does not exist for any \( \text{SNR} \) as explained in Section IV-C. However, for \( n_r \geq 2 \), the bounds do not exist for any \( \text{SNR} \) as explained in Section IV-C. The solutions to (22) and (23) lead to good upper bounds for any \( n_r \). Moreover, accurate bounds could be realised using these solutions even at high SNR where \( \beta = 0 \).

The SISO channel capacity shown by Abou-Faycal [5] with a discrete input is included in Fig. 6. It is slightly higher than \( C_{\text{sup}} \) at any SNR. It is clear that \( \beta = 0 \) provides the lowest upper bound as predicted in Section IV-C. However, for \( n_r \geq 2 \), the bounds do not exist for any \( \text{SNR} \) as explained in Section IV-C. The solutions to (22) and (23) lead to good upper bounds for any \( n_r \). Moreover, accurate bounds could be realised using these solutions even at high SNR where \( \beta = 0 \).

V. NUMERICAL RESULTS

For \( n_r = 1 \), the numerical results obtained with \( \beta = \Psi(n_r) + \log 2 \) in Fig. 6 is slightly higher than \( C_{\text{sup}} \) at any SNR. It is clear that \( \beta = 0 \) provides the lowest upper bound as predicted in Section IV-C. However, for \( n_r \geq 2 \), the bounds do not exist for any \( \text{SNR} \) as explained in Section IV-C. The solutions to (22) and (23) lead to good upper bounds for any \( n_r \). Moreover, accurate bounds could be realised using these solutions even at high SNR where \( \beta = 0 \).
VI. Conclusions

In this paper, we investigated the capacity of uncorrelated MIMO channels when neither the receiver nor the transmitter has the knowledge of CSI except its fading statistics. Capacity supremum is given in two cases, where one predicts accurate values for any number of receive antennas. The main findings of this paper is the capacity supremum for the non-coherent uncorrelated Rayleigh fading MIMO channel with no CSI for a given number of receiver antennas at any SNR.

We have shown the asymptotic behavior of the capacity supremum at high SNR. Furthermore, it is compared with the capacity shown in the literature having large number of receive antennas. The results of this paper can be used as an upper bound to the non-coherent uncorrelated Rayleigh fading MIMO channel with any input distribution.

The input distribution for the capacity upper bound is proven to be non-continuous. Hence the optimisation problem is posed for a discrete input with \( N \) mass points. Capacity obtained from the discrete input with optimal number of mass points, probabilities and their locations will be a lower bound to the capacity supremum derived in this paper. The capacity increase with increasing number of receive antennas is not promising at high SNR as in the coherent channel. Therefore, channel estimation becomes more important as the number of antenna elements at the receiver is increased.

VII. Appendix

A. Derivation of the constraint in (45)

We use (4) in LHS of (8c) to write

\[
\int_0^\infty p_Y(y) \log y dy = \int_0^\infty \int_0^\infty p_X(x) \left\{ \frac{\beta^{2n_r-1}}{2n_r-1} \right\} \exp \left\{ -\beta \frac{u^2}{2(1+x^2)^{n_r}} \right\} (\log y) dy dx.
\]

(44)

Using the integral identity [16]:

\[
\int_0^\infty x^t e^{-ux^2} \log x dx = \frac{u^{-\frac{t}{2}}}{4\sqrt{\pi}} \Gamma \left( t + \frac{1}{2} \right) \left[ \Psi \left( t + \frac{1}{2} \right) - \log u \right],
\]

(45)

we simplify (44) as

\[
\int_0^\infty p_Y(y) \log y dy = \frac{1}{2} \int_0^\infty p_X(x) \left[ \Psi(n_r) + \log 2(1+x^2) \right] dx
\]

\[
= \frac{1}{2} \beta + \Psi(n_r) + \log 2. \quad (46)
\]

This constraint is used in addition to the constrained input power in calculating the capacity supremum where \( \beta = E_x \{ \log(1+x^2) \} \).

VIII. Acknowledgements

National ICT Australia (NICTA) is funded through the Australian Government’s Backing Australia’s Ability Initiative, in part through the Australian Research Council. The authors would like to thank Michael Williams for helpful discussions.

References

[1] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” Tech. Rep., AT&T Bell Labs, 1995.
[2] W. Cao, W. U. Sjunjun, Z. Linrang, and T. Xiaoyan, “Capacity evaluation of MIMO systems by Monte-Carlo methods,” IEEE International Conf. on Neural Networks and Signal Processing, pp. 1464–1466, Dec. 2003.
[3] M. A. Khalighi, J. M. Brossier, G. Jourdain, and K. Raoof, “Water filling capacity of Rayleigh MIMO channels,” IEEE Journal on Selected Areas in Communications, vol. 01, pp. 159–138, Oct. 2001.
[4] Y. Liang and V. V. Veeravalli, “Capacity of noncoherent time selective Rayleigh fading channels,” IEEE Trans. on Info. Theory, vol. 50, no. 12, pp. 3095–3110, Dec. 2004.
[5] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, “The capacity of discrete-time memoryless Rayleigh-fading channels,” IEEE Trans. on Info. Theory, vol. 47, no. 4, pp. 1290–1301, May 2001.
[6] G. Taricco and M. Elia, “Capacity of fading channel with no side information,” IEEE Electronics Letters, vol. 33, no. 16, pp. 1368–1370, July 1997.
[7] J. Huang and S. Meyn, “Characterisation and computation of optimal distributions for channel coding,” in Proc. IEEE Trans. on Info. Theory, 2005.
[8] G. Smith, “The information capacity of amplitude and variance-constrained scalar Gaussian channels,” Information and Control, vol. 18, pp. 203–219, 1971.
[9] T. L. Marzetta and B. M. Hochwald, “Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading,” IEEE Trans. on Info. Theory, vol. 45, no. 01, pp. 139–157, Jan. 1999.
[10] L. Zheng and D. N. C. Tse, “Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel,” IEEE Trans. on Info. Theory, vol. 48, no. 02, pp. 359–383, Feb. 2002.
[11] A. Lapidoth and S. Moser, “On the fading number of multi-antenna systems over flat fading channels with memory and incomplete side information,” IEEE Int. Symp. on Info. Theory, vol. 01, pp. 478–478, June-July 2002.
[12] G. Foschini and M. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” Wireless Personal Communications, vol. 06, pp. 311–335, Mar. 1998.
[13] A. J. Goldsmith and P. P. Varaiya, “Capacity of fading channels with channel side information,” IEEE Trans. on Info. Theory, vol. 43, no. 06, pp. 1986–1992, Nov. 1997.
[14] R. G. Gallager, Information theory and reliable communication, John Wiley and Sons, USA, 1968.